A
TEXT-BOOK OF PHYSICS

A
TEXT-BOOK OF PHYSICS

BY

W. WATSON, A.R.C.S., B.Sc. (London)

ASSISTANT PROFESSOR OF PHYSICS AT THE ROYAL COLLEGE OF SCIENCE, LONDON

LONGMANS, GREEN, AND CO.

39 PATERNOSTER ROW, LONDON

NEW YORK AND BOMBAY

1899

PREFACE

THE following pages are primarily intended to form a text-book suitable for a student who is already familiar with the very elements of the subject. Nevertheless it is hoped that the references to the most elementary parts of the subject are sufficiently full to prevent the necessity of the reader having to refer to any other book. It has been the aim throughout to make the demonstrations of the various propositions considered as elementary as possible. Thus a knowledge of the elements of algebra and Euclid and of the meaning of the trigonometrical ratios is all that is assumed. Those sections which, on account of their difficulty or less importance as far as the sequence of the subject-matter is concerned, may well be omitted on a first reading, have been marked with an asterisk.

The settling of the order in which the various branches should be studied is a matter of some difficulty. Thus the strictly logical order, or at any rate the order which is most suitable from the standpoint of the nature of the phenomena dealt with, is often unsuitable in an elementary text-book. In such a work it is of the utmost importance that very little, if anything, should be taken for granted on account of the proof being postponed. The necessity for adopting an arrangement in which everything taken for granted in any section has been proved in the preceding sections, has been forced on me during my teaching work, at any rate for Elementary students. Thus in the following pages, in deciding on the order in which the subjects are dealt with, the question of the most convenient sequence from the point of exposition has been considered of paramount importance.

As no text-book can take the place of experimentally illustrated lectures and of practical work in the laboratory, no attempt has been made to describe experimental illustrations of the various phenomena. For the same reason the figures are entirely diagrammatic in character, and are not intended as *pictures* of apparatus, &c.; the object of the figures being to elucidate the text and not to take the place of the actual apparatus. In the preparation of the cuts I have received great assistance from Miss M. Reeks, who has taken great pains with the drawings, and in many cases has succeeded in making a clear and instructive diagram from very sketchy materials.

In conclusion I have to thank my colleague, Mr. S. W. J. Smith, M.A., for kindly looking through the section dealing with electrolysis, and for many valuable suggestions.

CONTENTS

BOOK I

MECHANICS AND PROPERTIES OF MATTER

PART I.—INTRODUCTORY

CHAPTER I

MATTER AND ENERGY

CHAPTER II

PHYSICAL QUANTITIES AND MEASUREMENTS

CHAPTER III

MEASUREMENT OF LENGTH

PART II.—KINEMATICS

CHAPTER IV

POSITION

CHAPTER V

MOTION OF TRANSLATION

CHAPTER VI

MOTION OF A RIGID BODY

CHAPTER VII

PERIODIC MOTION

PART III.—DYNAMICS

CHAPTER VIII

NEWTON'S LAWS OF MOTION

CHAPTER IX

EQUILIBRIUM OF FORCES

CHAPTER X

WORK AND ENERGY

CHAPTER XI

MACHINES

CHAPTER XII

FRICTION

CHAPTER XIII

GRAVITATION

CHAPTER XIV

THE PENDULUM

PART IV.—PROPERTIES OF MATTER

CHAPTER XV

PROPERTIES OF MATTER

CHAPTER XVI

PROPERTIES OF GASES

CHAPTER XVII

PROPERTIES OF LIQUIDS

CHAPTER XVIII

MOLECULAR PHENOMENA IN LIQUIDS

CHAPTER XIX

PROPERTIES OF SOLIDS

BOOK II

HEAT

CHAPTER I

THERMOMETRY AND EXPANSION BY HEAT

CHAPTER II

CALORIMETRY

CHAPTER III

CHANGE OF STATE

CHAPTER IV

CONDUCTION OF HEAT

CHAPTER V

RADIANT HEAT

CHAPTER VI

THE MECHANICAL THEORY OF HEAT

BOOK III

WAVE-MOTION AND SOUND

PART I.—WAVE-MOTION

CHAPTER I

WAVE-MOTION AND WATER WAVES

PART II.—SOUND

CHAPTER II

PRODUCTION AND PROPAGATION OF SOUND

CHAPTER III

PITCH—MUSICAL SCALE

CHAPTER IV

REFLECTION, REFRACTION, AND INTERFERENCE

CHAPTER V

VIBRATIONS OF STRINGS, RODS, PLATES, AND COLUMNS OF GAS

CHAPTER VI

SUPPLY OF ENERGY TO A SOUNDING BODY—RESONANCE

CHAPTER VII

AUDITION, COMBINATION TONES, CONSONANCE, AND VOCAL SOUNDS

BOOK IV

LIGHT

CHAPTER I

RECTILINEAR PROPAGATION—REFLECTION

CHAPTER II

REFRACTION

CHAPTER III

LENSES—MEASUREMENT OF REFRACTIVE INDEX

CHAPTER IV

PHOTOMETRY

CHAPTER V

VELOCITY OF LIGHT

CHAPTER VI

DISPERSION

CHAPTER VII

INTERFERENCE

CHAPTER VIII

EMISSION AND ABSORPTION OF LIGHT

CHAPTER IX

COLOUR SENSATIONS

CHAPTER X

POLARISATION AND DOUBLE REFRACTION

BOOK V

MAGNETISM AND ELECTRICITY

PART I.—MAGNETISM

CHAPTER I

MAGNETS AND MAGNETIC FIELDS

CHAPTER II

TERRESTRIAL MAGNETISM

PART II.—ELECTRO-STATICS

CHAPTER III

ELECTRO-STATIC ATTRACTION AND REPULSION—COULOMB'S LAW

CHAPTER IV

THE ELECTRICAL FIELD

CHAPTER V

CAPACITY—ELECTRICAL ENERGY

CHAPTER VI

ELECTROMETERS AND ELECTRICAL MACHINES

PART III.—ELECTRO-KINEMATICS

CHAPTER VII

THE ELECTRIC CURRENT

CHAPTER VIII

RESISTANCE

CHAPTER IX

JOULE'S LAW

PART IV.—THERMO-ELECTRICITY

CHAPTER X

THERMO-ELECTRICITY

PART V.—MAGNETIC INDUCTION

CHAPTER XI

MAGNETIC INDUCTION

PART VI.—ELECTRO-MAGNETISM

CHAPTER XII

FORCES ACTING ON CONDUCTORS CONVEYING CURRENTS

PART VII.—ELECTRO-MAGNETIC INDUCTION

CHAPTER XIII

INDUCED CURRENTS

CHAPTER XIV

ELECTRO-MAGNETIC MACHINES

PART VIII.—ELECTROLYSIS, ELECTROMOTIVE FORCE OF CELLS, AND PASSAGE OF ELECTRICITY THROUGH GASES

CHAPTER XV

ELECTROLYSIS

CHAPTER XVI

CONTACT E.M.F. AND THE VOLTAIC CELL

CHAPTER XVII

ENERGETICS OF THE VOLTAIC CELL

CHAPTER XVIII

PASSAGE OF ELECTRICITY THROUGH GASES

PART IX.—MAXWELL'S ELECTRO-MAGNETIC THEORY

CHAPTER XIX

TRANSFERENCE OF ELECTRO-MAGNETIC ENERGY AND MAXWELL'S ELECTRO-MAGNETIC THEORY OF LIGHT

CHAPTER XX

ELECTRICAL OSCILLATIONS

ERRATUM

Page 705, line 6 from end, *for* "Moisson" *read* "Moissan."

"Those sections which may well be omitted on a first reading are marked with an asterisk."

BOOK I

MECHANICS AND PROPERTIES OF MATTER

PART I—INTRODUCTORY

CHAPTER I

MATTER AND ENERGY

1. Province of Physics.—As a result of the observations and experiments made during many generations we are led to make certain assumptions or axioms which state that the physical universe has an objective existence, and that we are made acquainted with it solely by means of our senses. If further we give the name *thing* to that with the objective existence of which we are acquainted by our senses, then it follows that in the physical universe there are only two classes of things; to these the names *Matter* and *Energy* are given. Time and space, and many other quantities, such as Number, Velocity, Position, Temperature, &c., are not things.

It will probably be allowed at once that every form of matter, *i.e.* a stone, a drop of water, the air, &c., has objective existence; the most powerful argument in favour of this belief being the fact that all experiments have shown that under no circumstance whatever can we alter the quantity of matter. This result of experience, which is a fundamental assumption in all quantitative chemical experiments, is generally referred to as the *Conservation of Matter.*

The statement that energy has an objective existence is, however, one which is not so readily accepted: in fact its acceptance by scientific men only dates back a comparatively short time. Experiments, with which we shall deal later on, have however shown that energy, like matter, is indestructible and uncreatable by man. The objective existence is, as Professor Tait has pointed out, virtually admitted in a curious way by its being advertised for sale, it being quite common in manufacturing centres to see the notice "Spare Power to Let." Again, water under a great

pressure is supplied for the purpose of working hydraulic lifts, &c., and since the price paid for a given quantity of water is in these circumstances much higher than that for which the same quantity of water would be obtained at such pressures as are found in the ordinary supply mains, we infer that the purchaser thinks he is buying some "thing" besides the matter of which the water is composed.

From the foregoing considerations we are led to define Physics in its most general aspect as a discussion of the properties of matter and energy. It is, however, usual to restrict somewhat the definition so as to exclude the discussion of those properties of matter which depend simply on the nature of the different forms of matter (Chemistry), as also the properties of matter and energy as related to living things (Biology). The line of demarcation separating Physics and Chemistry has never been very clear, and of late years has practically vanished.

2. Matter.—Of the numerous definitions of matter which have from time to time been given, we may at present adopt the following: Matter is that which can occupy space. This definition does not attempt to state what matter *is*, it only gives us a working definition, which in the present state of our knowledge as to the ultimate structure of matter is all that can be done.

We may speak of a limited portion of matter as a body, and of matter of a certain definite kind as a substance. Thus water, sugar, air, lead, are all matter, since they all occupy space or have dimensions. Since each of these things is a special kind of matter possessing distinct properties, they each form a distinct substance. A drop of water, a lump of sugar, the air enclosed in a given vessel, is each an example of a body.

3. Energy.—Energy may be defined as the capacity of doing work, where by work we mean the act of producing a change of the state of matter in opposition to resistance, which opposes any such change. The real meaning of this definition will be made clearer when we come to consider the various forms in which energy can exist.

4. General Maxim of Physical Science.—There is a maxim to the effect that the same cause will always produce the same effects, which is at the foundation of all our investigations in Physical Science. Since no event ever happens more than once, it is evident that the causes and effects spoken of above cannot be the same in all respects. What is meant is that if the causes only differ as regards the absolute time and place at which the event we are considering occurs, so the effects will also only differ as regards the absolute time and place. In order to meet this defect in the maxim, Maxwell has proposed to substitute the following: "The difference between one event and another does not depend on the mere difference of the times or the places at which they occur, but only on differences in the nature, configuration, or motion of the bodies concerned.

It follows that if a certain event has happened under a certain definite set of conditions, then if at any time exactly the same conditions again arise, a similar event must necessarily follow.

The belief in the truth of this maxim is at the foundation of all experiments, for an experiment is simply the artificial arrangement of certain causes, so that we may determine how, when one or more of the causes is inoperative, the event differs from that observed when all the causes ordinarily present are effective. If, then, by experiment we find that certain causes are allied to certain effects, we feel sure that the same causes and the same effects will always be allied; while if in any experiment the effect observed varies when we keep constant all the causes that, as far as we know, are operative, then we may at once assume that there is some other cause besides those we have taken into account which is varying and causing the variation in the effects; and it is by investigating such causes that our knowledge of nature is gradually extended.

CHAPTER II

PHYSICAL QUANTITIES AND MEASUREMENTS

5. Physical Magnitudes.—Although in some cases we may not be able to measure it with any great accuracy, every physical quantity has a certain definite magnitude. Whatever the nature of the physical quantity may be, we employ to measure its magnitude a certain fixed amount of the *same* kind of physical quantity, which we call the *unit* of that particular quantity. The given quantity is then said to be equal to so many times the unit.

Thus, in order to measure the magnitude of a given length, we take as our unit some standard length, say the yard, and then find how many times this length will go into the given length. Say it goes x times, where x may be a whole number or a proper or improper fraction, then the given length is said to be x yards. We see, therefore, that the complete statement of the result of a measurement of a physical quantity consists of two parts; first, a pure number, called the numeric, which states the number of times the unit is contained in the given quantity; and, second, the name of the unit which has been employed. Every statement of the magnitude of a physical quantity must consist of these two parts, or it will be ambiguous. Thus if we were to say that a certain length was three, it would be uncertain whether we meant three inches, or three feet, or three miles, &c.

6. Units.—Since the magnitude of every physical quantity has to be measured in terms of a unit of its own kind, it follows that there will be as many units as there are different kinds of physical quantities to be measured.

As great inconvenience would be caused if different people used in their measurements different units, the magnitude of the unit has in most cases, either by usage or by law, been agreed upon. Such a unit is generally called a standard unit.

7. Fundamental and Derived Units.—The magnitude of the unit chosen in every case may, if we like, be quite arbitrary, and in fact until a comparatively recent time this was so. The advances of physical science have, however, shown that there are certain relations which exist between different kinds of physical magnitudes, and that by selecting the units in a certain number of cases it is possible, by making use of these relations, to fix the magnitude of the units for the rest of the

physical quantities. The units which are thus chosen as the basis for our system of units are called *fundamental units*, while those units, for the determination of the magnitude of which we make use of the relations which exist between the physical quantity in question and the fundamental units, are called *derived units*.

The physical quantities which are most commonly employed as fundamental units are those of length, mass, and time, although energy or force is sometimes employed as a fundamental unit in place of mass. In either case it is found that, with a few exceptions which are probably caused by our ignorance of the true nature of the phenomena considered, and which will be referred to later, it is possible to fix the magnitude of the unit to be employed in the case of all other physical quantities when we have fixed the value of these three fundamental units.

As examples of fundamental units we may take the yard, which is one of the British standard units of length, or the second, which is the unit of time. The gallon and pint, which are used as units of volume, have no connection with the unit of length. If, however, we take as our unit the volume of a cube, of which each edge is of unit length, then there is a direct connection between the unit of volume, which is in this case a derived unit, and the unit of length, a fundamental unit. Again, the velocity with which light traverses interstellar space is sometimes taken as the unit of velocity; this unit has no direct connection with the units of length and time. If the unit velocity, however, is defined as such that a body travelling with this velocity passes over the unit of length in the unit of time, then we have a direct connection between these three units, so that being given the magnitude of the two fundamental units of time and of length, we can at once say what is the unit of velocity.

8. Absolute Systems of Units.—A system of units in which certain units are chosen as fundamental, and all the others are derived units connected with these by fixed physical relations, is called an *absolute system*; measurements made in terms of these units being said to be in absolute units. The word absolute is sometimes used in a slightly different sense, *i.e.* as an antithesis to relative. For example, if a velocity is measured by comparing it with some known velocity, we are said to make a relative measurement. If, however, the velocity is measured by determining the length passed over by the body in the unit of time—*i.e.* if the quantities we actually measure are the fundamental quantities, length and time—we are said to make an absolute measurement. It must be carefully borne in mind that the word absolute has here no reference whatever to the accuracy or inaccuracy of the observations.

The term "absolute system of units" was first introduced by Gauss in 1832, in connection with his measurements of the strength of the earth's magnetic field at Göttingen. Instead of measuring, as had been done up to that date, this quantity in terms of the strength of the earth's field at some fixed place (such as London) taken as the unit, Gauss expressed

it in terms of the units of length, mass, and time, and thus the value of the unit did not change as the strength of the earth's field changed at the standard place, as was the case before.

There are several absolute systems of units possible according (1) to what we take as the fundamental units, (2) to the magnitude we adopt for the fundamental units chosen, (3) to the physical relation we employ for obtaining the derived units from the fundamental units. Thus we may take as our fundamental units those of length, mass, and time, or of length, force, and time, or of length, energy, and time, or length, mass, and force, &c. &c. Again, we may take as our unit of length the yard, the inch, the mile, or the metre. Finally, we might define the unit of volume as the volume of a cube, each edge of which is of unit length, or as the volume of the sphere whose radius is of unit length.

With the exception of the electrical units, it is with reference to the first two of these three possible modes of variation that all practical absolute systems differ amongst themselves. By far the most usual system in all physical investigations is that in which the fundamental units are those of length, mass, and time, and in which the unit of length is the centimetre, the unit of mass the gram, and the unit of time the second. This system is referred to as the *c.g.s.* (centimetre, gram, second) system. This is the system that will be almost exclusively used in this volume, though occasionally, where there are other units in common use, they will be referred to, in order to familiarise the reader with the actual magnitude of the *c.g.s.* units.

An absolute system, which till quite lately was employed in all English Observatories, and is in fact still employed in some, is that in which the unit of length is the foot, the unit of mass the grain, and the unit of time the second. Again, the foot, the pound, and the second are sometimes (chiefly, let it be said, in text-books on mechanics, and in examination papers) used as the fundamental units.

A more important system of absolute units is that in which the fundamental units are those of length, force, and time, for this system, which will be referred to later as the gravitational system, is almost exclusively used by engineers (*at any rate in this country*).

Finally, there is the system proposed by Ostwald, in which the fundamental units are those of length, energy, and time.

9. Dimensions of Derived Units.—The relation by means of which we derive the magnitude of the unit of any quantity, in terms of the fundamental units, is indicated by what is called the *dimensions* of the unit in question. The easiest way to see how this is done will be to consider some simple examples.

As has been stated in § 5, the record of any quantity, say a length, must consist of two parts, a pure number and a term giving the name of the unit employed. Thus we may indicate any length by the symbol $l[L]$, where l represents the numerical part of the expression, *i.e.* the

number of times the unit is contained in the given length, and $[L]$ is used as a symbol to represent the unit of length employed. In the same way $[M]$ and $[T]$ represent the units of mass and time respectively. The unit of area could be indicated by the symbol $[A]$. If, however, we use an absolute system in which the relation between the unit of area and that of length is that the unit of area is the area of a square of which the sides are each of unit length, then we may write the symbol for this unit $[L.L]$ or $[L^2]$, since the area of a square is numerically equal to the square of the length of one of the sides. It is evident that if the unit of length be increased from $[L]$ to $[L+L']$, then the unit of area will be increased from $[L^2]$ to $[(L+L')^2]$. Hence we may say that the statement that the unit of area can be represented by $[L^2]$ at once tells us how a change in the fundamental unit of length affects the derived unit of area. Since the unit of area would not change if the units of mass and time were changed, the relation between the unit of area and these units may be represented by $[M^0]$ and $[T^0]$. Hence, collecting these three symbolical statements into one, we may say that the unit of area $[A]=[L^2]\,[M^0]\,[T^0]$ or $[A]=[L^2.M^0.T^0.]$, which is interpreted as meaning that in the absolute system we are using, the unit of area varies as the second power, or as the square, of the unit of length, but does not vary with the units of mass and time. We therefore say that the *dimensions* of the unit of area with reference to length, mass, and time are 2, 0, 0, or more fully, so as to leave no doubt as to the order in which we are referring to the units, $[L^2.M^0.T^0.]$. It will be at once evident that the dimension of the unit of volume are $[L^3.M^0.T^0.]$.

To take another example, consider the unit of velocity, which in our absolute system is defined as such that unit space is passed over in unit time. If we double the unit of length, keeping the unit of time constant, we shall evidently require the body to move over twice the distance in the unit of time, *i.e.* we shall double the unit of velocity; therefore the unit of velocity $[V]$ has the dimensions 1 with reference to the unit of length. Again, if we double the unit of time, we allow the body twice as long to cover the unit of length, supposed to remain constant, and therefore we halve the unit of velocity. Hence the dimensions of velocity with reference to time are -1. This is an abbreviation for

$$\left[\frac{1}{T}\right] \text{ or } [T^{-1}].$$

As the unit of velocity does not depend on the unit of mass, its dimensions must therefore be $[L^1.M^0.T^{-1}.]$.

10. Dimensional Equations.—Equations such as

$$[A]=[L^2.M^0.T^0.] \text{ or } [V]=[L^1.M^0.T^{-1}.],$$

which tell us the relation between the derived unit and the fundamental units of a system, are called *dimensional equations*. They are of utility in

two ways—(1) as affording a means by which we can convert the magnitude of any physical quantity expressed in terms of the units belonging to one absolute system into those of any other absolute system; (2) they afford a check on the accuracy of the line of reasoning by means of which we have deduced an equation connecting any physical quantities. Since it is impossible to compare two physical quantities which are not of the same kind, it follows that the dimensions of the two sides of any equation connecting physical quantities must be the same. Thus, suppose we had come to the conclusion that the volume c of water, which passes any point of a river during the time t, was given in terms of the area of cross-section a of the river, and the velocity v by the equation

$$c=v^2at \quad . \quad . \quad . \quad (1).$$

Substituting for c, v, a, t, the full expressions in which the values of the units appear, we get

$$c[C]=v^2[V^2]\times a[A]\times t[T].$$

Then when each of the quantities c, v, a, and t are unity, we get the dimensional equation

$$[C]=[V^2][A][T].$$

Substituting on each side of this equation in terms of the fundamental units, we get

$$[L^3]=[L^2T^{-2}][L^2][T]$$
$$=[L^4T^{-1}].$$

Here the dimensions on the two sides are different, and hence we conclude that the assumption made in equation (1) is incorrect. As a matter of fact, the correct equation is

$$c=vat,$$

and this gives the dimensional equation

$$[L^3]=[L^1T^{-1}][L^2][T]=[L^3],$$

in which the dimensions on the two sides are the same. As an example of the use of dimensions for changing from one system of units to another, suppose it is required to convert a velocity of x miles per hour into feet per second. Here we have two units of length, the mile and the foot; let us represent them by $[L]$ and $[L']$ respectively; in the same way take $[T]$ and $[T']$ to represent an hour and a second. If y is the numeric which expresses the velocity in feet per second, we must have, since the actual value of the velocity remains the same whatever units we may employ in which to measure it—

$$x[LT^{-1}]=y[L'.T'^{-1}.]$$

$$\therefore \qquad y=x\left[\frac{L}{L'}\cdot\frac{T'}{T}\right].$$

This shows that to obtain y, the value of the velocity expressed in feet

per second, we must multiply the number which expresses the velocity in terms of miles and hours by the ratio of the mile to the foot $\left(\frac{L}{L'}\right)$ and by the ratio of the second to the hour $\left(\frac{T'}{T}\right)$.

Hence $$y = x \times \frac{5280}{1} \times \frac{1}{3600}.$$

The further discussion of dimensions will be postponed till later, but in order that the reader may gradually familiarise himself with the dimensions of different physical quantities, the dimensions of each quantity will be given at the place where this quantity is under discussion.

11. Units of Length.—There are in Great Britain two standard units of length—the yard and the metre. The yard is defined by Act of Parliament[1] as follows: "The straight line or distance between the centres of the transverse lines in the two gold plugs in the bronze bar deposited in the Office of the Exchequer[2] shall be the genuine standard yard at 62° F., and if lost it shall be replaced by means of its copies." Copies of the standard yard are deposited at the Royal Mint, the Royal Society of London, the Royal Observatory at Greenwich, and the Houses of Parliament.

The second standard unit of length in Great Britain[3] is the metre. The metre owes its origin to a law of the French Republic,[4] which enacted that the unit of length should be one ten-millionth $\left(\frac{1}{10^7}\right)$ of the distance between the North Pole and the Equator, measured over the surface of the earth along the meridian passing through Paris. The measurement of the arc of this meridian between Barcelona and Dunkirk was carried out by Delambre and Méchain, and from their results Borda constructed the standard metre to fulfil the above definition. The metre is now, however, not defined as the ten-millionth of the quadrant of the meridian, but as the distance between the ends of Borda's platinum rod at a temperature of 0° C. If this were not so, each time a more accurate measurement of the earth's dimensions was made, all the copies of the metre in general use would have to be altered. Since the unit of length is a fundamental unit, we are able to keep it unaltered. If, however, it had not been a fundamental unit, it would have been necessary to alter its value each time a more accurate determination was made of it, in terms of the fundamental units. The inconvenience which may thus arise will be noticed when we come to the consideration of the electrical

[1] 18 & 19 Vict. c. 72, July 30, 1855.

[2] In accordance with the Weights and Measures Act of 1878, the British standards are now preserved at the Standards Office of the Board of Trade.

[3] The use of the metre as a unit of length was legalised by Act of Parliament in 1897.

[4] Loi du 18 Germinal, an iii. (April 7, 1795).

units. According to the best modern measurements, the length of the earth's quadrant is 10,000,880 metres.

The relation between the metre and the yard is

$$1 \text{ metre} = 39.37079 \text{ inches} = 1.093633 \text{ yards},$$
$$1 \text{ yard} = 0.9143935 \text{ metres}.$$

Although, looked at simply as a standard unit of length, the metre is not in any way preferable to the yard, yet, on account of the fact that the multiples and sub-multiples of the metre are all decimally connected with the metre, it is very much more convenient to use the metre as the standard.

A metre (m.) is divided into ten decimetres (dm.), a decimetre into ten centimetres (cm.), and a centimetre into ten millimetres (mm.). The only multiple of the metre practically employed is the kilometre, which is equal to one thousand metres, and is the unit adopted on the Continent for stating such distances as we should state in miles. The relation between the mile and the kilometre is 1 kilometre = 0.6214 mile, or 1 mile = 1.6093 kilometres.

For scientific purposes the centimetre is, in accordance with the recommendations of a Committee of the British Association, almost exclusively used as the unit of length.

For measuring very small lengths the following fractions of a millimetre are often employed:—The micron, equal to a thousandth of a millimetre (0.001 mm.), and the micromillimetre, equal to a millionth of a millimetre (0.000001 mm. or 10^{-6} mm.). The micron is often indicated by the symbol μ, and a micromillimetre by $\mu\mu$.

It is sometimes useful to remember that the diameter of a halfpenny is 1 inch, and that of a French dixcentime piece is 3 centimetres; also 1 inch = 25.4 millimetres (very nearly).

The following tables will be found of use in converting from the British to the metric system, and *vice versa*:—

British to Metric.		
Inches = Centimetres.	Feet = Centimetres.	Yards = Metres.
1 = 2.5400	1 = 30.479	1 = 0.91438
2 = 5.0800	2 = 60.959	2 = 1.82877
3 = 7.6199	3 = 91.438	3 = 2.74315
4 = 10.1598	4 = 121.918	4 = 3.65753
5 = 12.6998	5 = 152.397	5 = 4.57192
6 = 15.2397	6 = 182.876	6 = 5.48630
7 = 17.7797	7 = 213.356	7 = 6.40068
8 = 20.3196	8 = 243.835	8 = 7.31507
9 = 22.8596	9 = 274.315	9 = 8.22945

Metric to British.		
Centimetres = Inches.	Centimetres = Feet.	Metres = Yards.
1 = .39371	1 = .032809	1 = 1.09363
2 = .78742	2 = .065618	2 = 2.18727
3 = 1.18112	3 = .098427	3 = 3.28090
4 = 1.57483	4 = .131236	4 = 4.37453
5 = 1.96854	5 = .164045	5 = 5.46817
6 = 2.36225	6 = .196854	6 = 6.56180
7 = 2.75596	7 = .229663	7 = 7.65543
8 = 3.14966	8 = .262472	8 = 8.74906
9 = 3.54337	9 = .295281	9 = 9.84270

12. Units of Mass.—The standard units of mass in Great Britain are the pound, which is the unit in the British system, and the kilogram, which is the unit in the metric system.

The standard pound avoirdupois is the mass of a certain piece of platinum which is marked "P. S., 1844, 1 lb.," and is kept at the same place as the standard yard. The grain, which has been used as the unit of mass in the old British absolute system of units, is one seven-thousandth part of the pound.

The kilogram is the mass of a certain lump of platinum which is preserved at Paris, and is called the "Kilogram des Archives." The kilogram was originally constructed by Borda, to represent the mass of a cubic decimetre, that is 1000 c.c. of water at 4° C., the temperature of maximum density. More recent measurements have, however, shown that it does not exactly fulfil this definition (see § 146).

The *c.g.s.* unit of mass is one thousandth of the kilogram, and is called the gram. A thousandth of the gram is called the milligram. The following are the usual abbreviations used to represent the various units of mass—pound, lb.; kilogram, kilo.; gram, grm.; milligram, mgrm.

A kilogram is equal to 2.20462125 lbs. or 15432.34874 grains, and a pound is equal to 0.45359265 kilos.

The following tables will be found of use for converting from the British to the metric system, and *vice versa*:—

British to Metric.		
Pounds = Kilograms.	Ounces = Grams.	Grains = Milligrams.
1 = 0.45359	1 = 28.3495	1 = 64.79895
2 = 0.90719	2 = 56.6991	2 = 129.59790
3 = 1.36078	3 = 85.0486	3 = 194.39685
4 = 1.81437	4 = 113.3982	4 = 259.19580
5 = 2.26796	5 = 141.7477	5 = 323.99475
6 = 2.72156	6 = 170.0972	6 = 388.79370
7 = 3.17515	7 = 198.4468	7 = 453.59265
8 = 3.62874	8 = 226.7963	8 = 518.39160
9 = 4.08233	9 = 255.1459	9 = 583.19055

Metric to British.	
Kilograms = Pounds.	Grams = Grains.
1 = 2.20462	1 = 15.43235
2 = 4.40924	2 = 30.86470
3 = 6.61386	3 = 46.29705
4 = 8.81849	4 = 61.72939
5 = 11.02311	5 = 77.16174
6 = 13.22773	6 = 92.59409
7 = 15.43235	7 = 108.02644
8 = 17.63697	8 = 123.45879
9 = 19.84159	9 = 138.89114

13. Units of Time.—The scientific unit of time, both in the British and the metric system, is the mean solar second. The mean solar second is one 86,400th part of a mean solar day. The mean solar day is the average interval which elapses between successive transits of the sun across the meridian at any place during a whole year.

Owing to the excentricity of the earth's orbit, and the fact that the earth's axis is not perpendicular to the plane of the orbit, the interval between two successive transits varies during the year, so that the actual solar day is not the same as the mean solar day.

If a clock keeps mean time and agrees with solar time, that is time such as would be indicated by a sundial, when the sun appears in that portion of the heavens known as the first point of Aries, then the difference between the time of noon as indicated by this clock, and the time when the sun crosses the meridian on any day, is called the equation of time at noon for that day. The curve in Fig. 1 gives the equation of time for the year. When the curve is above the axis OX the equation of time is positive, that is, the time as shown by a mean-time clock will be ahead of the transit of the sun by the amount shown by the ordinate.

It will be seen that the equation of time is zero, that is, the time

as indicated by a mean-time clock and the sun will be the same on April 15, June 15, August 31, and December 24. On February 11 the mean-time clock is 14 minutes 29 seconds ahead of the sun, while on November 1 it is 16 minutes 20 seconds behind the sun.

The unit of time used in astronomy is the sidereal day. This represents the interval between two consecutive transits of one of the fixed stars across the meridian. Since the distance between the earth and any of the fixed stars is very great, compared even with the diameter of the earth's orbit, the line joining the earth to such a star remains always parallel to itself. Hence the sidereal day represents the time the earth takes to make one complete rotation about its axis. A sidereal day is equal to 23 hours 56 minutes 4.09 seconds of mean solar time.

The use of the rotation of the earth as a measurer of time is not without objection, for there can be no doubt that the mean solar day is gradually growing longer, due to the slowing down of the rotation of the earth. In order to remove this objection, it has been proposed to use the time of vibration of the atom of some element, such as sodium, as the unit of time, for under definite conditions it appears as if this time were quite fixed and unalterable.

14. Units of Angular Measurement.—The ordinary unit adopted for measuring angles is the degree: 90 degrees being equal to a right angle, so that 360 degrees correspond to a complete rotation. Each degree is divided into 60 minutes, and each minute into 60 seconds. Degrees, minutes, and seconds of arc are indicated by the symbols °, ′, and ″ respectively.

Another unit of angle, which is frequently employed, is called the *radian*, and is such that if an arc of the circumference of a circle is taken equal in length to the radius of the circle, then this arc will subtend an angle at the centre which is equal to one radian. When the radian is used as the unit, the angle is said to be measured in *circular measure*.

If we have an arc of a circle, of which the length is a, then the angle subtended at the centre of the circle by this arc is equal to a/r radians, where r is the radius of the circle. When $a=r$ this of course reduces to one radian according to the definition.

Since the length of the circumference of a circle of radius r is $2\pi r$, this arc will subtend $2\pi r/r$ or 2π radians at the centre. But the angle subtended at the centre by the whole circumference is 360°. Hence

$$
\begin{aligned}
2\pi \text{ radians} &= 360^\circ. \\
\therefore\ 1 \text{ radian} &= 360^\circ/2\pi. \\
&= 57^\circ.2958. \\
&= 57^\circ\ 17'\ 44''.88.
\end{aligned}
$$

Also $$1 \text{ degree} = 0.017453 \text{ radians}.$$

If AB (Fig. 2) is an arc of a circle of radius r, described about the point O as centre, then the angle θ, subtended by this arc at the centre, is equal to $\overline{AB}/r$ radians. If from B we draw BC perpendicular to OA, then the following trigonometrical relations hold good :[1]—

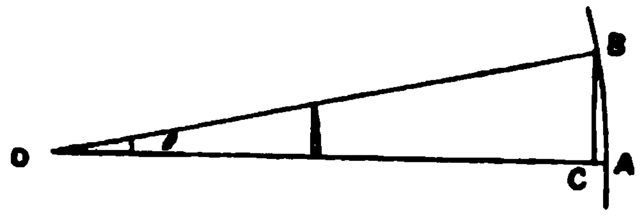


FIG. 2.

$$\theta \text{ (in circular measure)} = \overline{AB}/r.$$
$$\text{Sin } \theta = \overline{BC}/\overline{OB} = \overline{BC}/r.$$
$$\text{Cos } \theta = \overline{OC}/\overline{OB} = \overline{OC}/r.$$
$$\text{Tan } \theta = \overline{BC}/\overline{OC}.$$

Now, if the angle θ is very small, the length of the arc $\overline{AB}$ will be very nearly the same as that of the perpendicular $\overline{BC}$, while $\overline{OC}$ will be very nearly equal to $\overline{OB}$ or r. Hence, when θ is very small, the above relations reduce to—

$$\theta = \overline{AB}/r.$$
$$\text{Sin } \theta = \overline{AB}/r.$$
$$\text{Cos } \theta = r/r = 1.$$
$$\text{Tan } \theta = \overline{AB}/\overline{OA} = \overline{AB}/r.$$

Hence for small values of θ we have—

$$\text{Sin } \theta = \tan \theta = \theta,$$

where θ is measured in radians, and

$$\text{Cos } \theta = 1.$$

The closeness with which these relations are true for different small values of θ will be evident from the following table :—

θ In Degrees.	θ In Radians.	Sin θ.	Tan θ.	Cos θ.
$0°$	0	0	0	1
$0° 30'$	0.00873	0.00873	0.00873	0.99996
$1°$	0.01745	0.01745	0.01746	0.99985
$2°$	0.03491	0.03490	0.03492	0.99939
$3°$	0.05236	0.05234	0.05241	0.99863
$4°$	0.06981	0.06976	0.06993	0.99756
$5°$	0.08727	0.08716	0.08749	0.99619

[1] Readers unfamiliar with the elements of trigonometry may take these relations as *defining* the quantities—the sine of the angle θ (written sin θ), the cosine of θ (cos θ), and the tangent of θ (tan θ).

These relations will be found of considerable utility, for by their means we are able, whenever we are dealing with small angles, to considerably simplify many expressions involving these functions of the angle.

Since an angle is measured, in circular measure, by the ratio of the length of the arc (a) to the radius (r), we have, if $[\theta]$ is taken to represent the dimensions of the unit angle, the relative

$$\begin{aligned}[\theta] &= [L] / [L] \\ &= 1.\end{aligned}$$

Thus an angle has dimensions zero with reference to all the fundamental units. As the dimensions of any quantity cannot depend on the absolute value of the unit used to measure it, it follows that an angle, when measured in degrees, is also of zero dimensions with reference to the fundamental units of length, mass, and time.

CHAPTER III

MEASUREMENT OF LENGTH

15. Importance of Length Measurements.—In the measurement of nearly all physical quantities, what we actually *observe* is the ratio of some length to some other length. Thus when we measure the pressure of the atmosphere with a mercury barometer, what we really observe and measure is the length of a column of mercury; the same statement applies to the measurement of a temperature with a mercurial thermometer; so also, when we use a spring balance to measure a mass, it is the movement of a pointer along a scale that is observed. Hence we see the importance of being able to make accurate measurements of length.

With an ordinary scale divided into tenths of an inch it is possible, with a little care and practice, to measure by eye a length, which is not greater than that of the scale, to within one hundredth of an inch. This is done by mentally supposing each of the tenths of an inch subdivided into ten equal parts, *i.e.* into hundredths of an inch, and estimating by eye by how many of these imaginary hundredth of an inch divisions the length exceeds the nearest number of whole divisions. In the same way, with a scale divided into millimetres, it is possible to read to tenths of a millimetre. In order to attain to a degree of accuracy much greater than the above it is necessary to adopt some mechanical means of subdividing the divisions, for merely making the divisions of the scale nearer together does not advance matters much if we trust to our judgment and eye alone, even if a magnifying glass is used. Of such mechanical contrivances the most commonly employed are the vernier and the micrometer screw.

16. The Vernier.—Suppose AB, Fig. 3, is a scale divided into equal

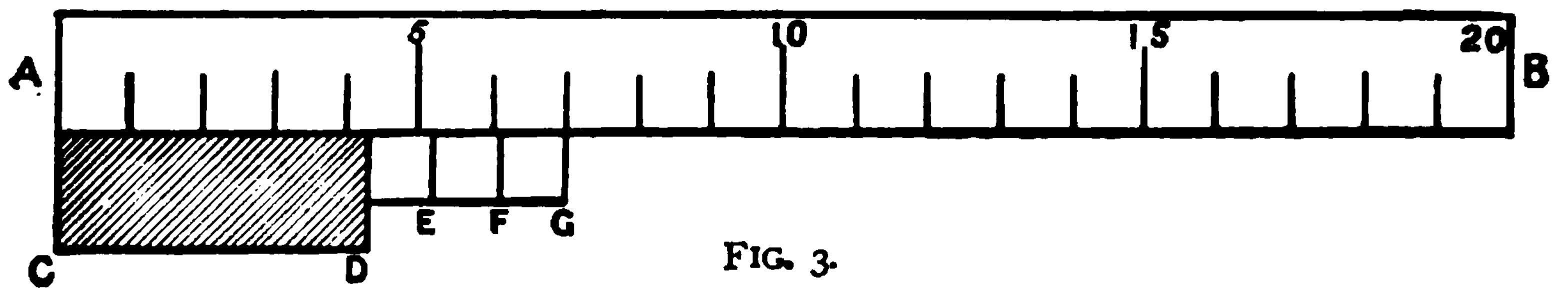


FIG. 3.

parts, and that the end D of some object (CD), the length of which is to be measured, lies between the fourth and fifth divisions, and that we

require to find the fraction (x) of a division by which the length of CD exceeds four divisions of the scale. If a block DE, of such a length that ten of these blocks placed end to end would be equal to nine divisions of the scale (*i.e.* the length of each block is $\frac{9}{10}$ or 0.9 of a division), is placed at the end of CD, then the end (E) of this block will evidently project beyond the fifth division by an amount $(x-\frac{1}{10})$ of a division, since DE is $\frac{9}{10}$ of a division. If a second block EF is placed at the end of the first, the end F will exceed the sixth division of the scale by an amount $(x-\frac{2}{10})$ of a division. In the same way the end of a third block would project beyond the seventh division of the scale by an amount $(x-\frac{3}{10})$ of a division, and so on. It will be seen, however, that the end G of the third block exactly coincides with the seventh division of the scale, so that the amount by which it projects is zero. Hence

$$x-\tfrac{3}{10}=0,$$

or

$$x=\tfrac{3}{10}.$$

That is, the length of $\overline{\text{CD}}$ is $4\frac{3}{10}$ or 4.3 divisions of the scale. We notice that if each of the blocks is $\frac{9}{10}$ths of a division in length, the object CD exceeds four divisions by as many tenths of a division as it is necessary to add blocks, till the end of the last block just coincides with one of the divisions of the scale.

It will be found that the above relation between the number of the blocks and the excess (x) is quite general, and it is utilised to mechanically subdivide the smallest divisions of a scale. Instead of having a number of separate blocks, it is more convenient to have a small auxiliary scale, called a vernier, which can slide along the edge of the chief scale, and is divided so that ten divisions of the vernier are equal to nine divisions of the scale. In this case we set the end of the vernier against the end of the object, and look along till we come to the division of the vernier which coincides with one of the divisions of the scale. In the case shown in Fig. 4, this occurs at the seventh division of the vernier, and hence the object is $4\frac{7}{10}$ or 4.7 divisions in length.

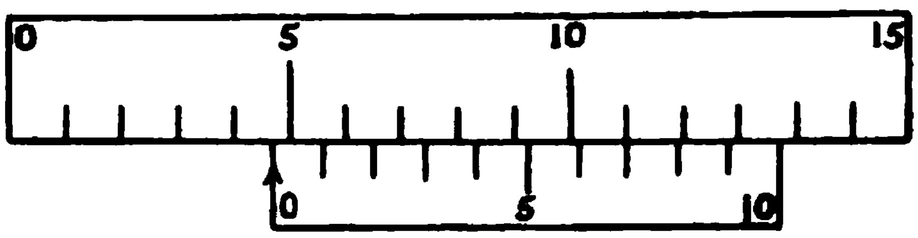


FIG. 4.

We may generalise and say that if m divisions of the vernier are equal in length to $m-1$ divisions of the scale, and coincidence occurs at the nth division of the vernier, then the reading of the vernier, *i.e.* the distance between the zero line of the vernier and the preceding division line on the scale, is $\frac{n}{m}$ths of a division of the scale. Thus if the scale is divided into millimetres, 20 divisions of the vernier being equal to 19 mm.,

and coincidence occurred at the 17th division of the vernier, the reading would be $\frac{17}{20}$ mm.

Verniers are sometimes constructed so that m divisions of the vernier are equal to $m+1$ divisions of the scale. For instance, suppose ten divisions of the vernier are equal to eleven divisions of the scale. Then, by an argument exactly similar to that adopted above, it is evident that if the nth division of the vernier, counting as before from the end next the object which is being measured, coincides with a division of the scale, we have

$$x+\frac{n}{10}=1 \text{ division,}$$

or

$$x=\frac{10-n}{10} \text{ of a division.}$$

But $10-n$ is the number of divisions of the vernier between the coincidence and the end remote from the object. Hence if the vernier is numbered in the reverse direction to that in which the scale is numbered, the reading on the vernier will give directly in tenths the fraction of a division. The advantage of this form of vernier is that it is a little more open, *i.e.* the divisions are further apart, than in the other form.

17. The Micrometer Screw.—If a screw is rotated through a complete turn its point will move, with reference to the nut, through a distance equal to the pitch of the screw, *i.e.* to the distance between two consecutive threads. By making the pitch of a screw small, and also attaching a drum-shaped head of considerable diameter which is divided into a number of equal parts, so that a fraction of a rotation can be read, it is possible to measure with great accuracy the distance moved over by the point of the screw. For example, in the Whitworth measuring machine the pitch of the screw is $\frac{1}{20}$ inch, so that the point of the screw advances $\frac{1}{20}$ of an inch for each whole turn. The head attached to the screw is divided into 500 equal parts. Hence one division on the head corresponds to a movement of the end of the screw of $\frac{1}{500}$ of $\frac{1}{20}$ or 1/10,000 inch.

18. The Screw-Gauge.—An example of a case where a micrometer screw is used to measure a length is afforded by the screw-gauge shown in Fig. 5. The object to be measured is placed between the end of the screw A and the block B, which is connected by a strong curved arm with the nut in which the screw works. The number of whole turns made by the screw is read by means of a scale, E, attached to the nut, which is gradually uncovered by the movement of the cap, G, attached to

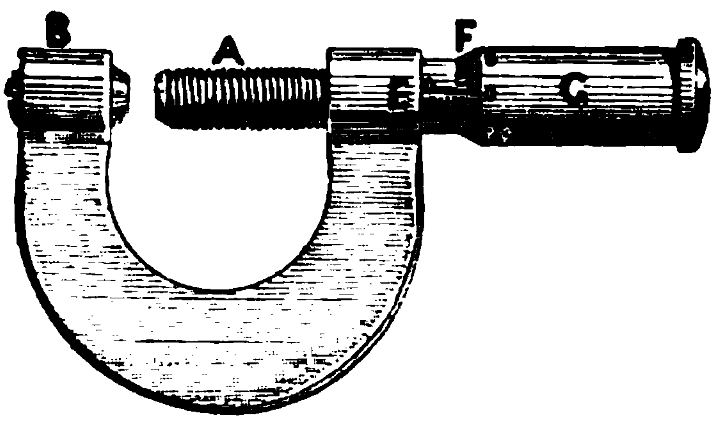


FIG. 5.

the screw. The fractions of a turn are read off on a scale, F, on the edge of this cap. The pitch of the screw ordinarily employed is 0.5 mm., and the edge of the cap is divided into 50 parts. Hence as turning the screw through a whole turn or 50 divisions advances the point A by 0.5 mm., one division on the scale F corresponds to a motion of the point of 1/50 of 0.5 mm. or 0.01 mm.

19. The Comparator.—For comparing together two very nearly equal lengths, as, for instance, a standard metre with a copy, an instrument called a comparator is used. The principle on which this instrument works will be seen from Fig. 6. Two stone pillars, A and B, which are firmly embedded in the ground, carry two microscopes, C and D. The cross wires of these microscopes, instead of being fixed, can be moved

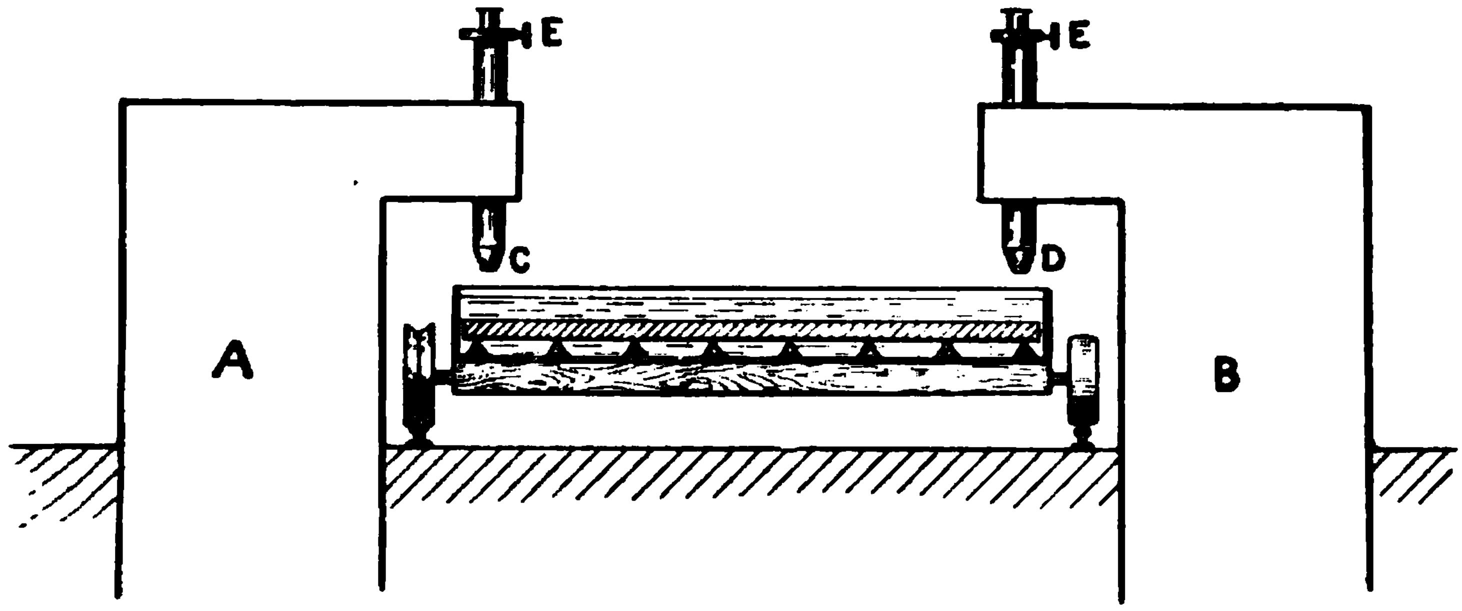


FIG. 6.

through a small distance in the direction of the line joining the pillars by means of micrometer screws, E, which have divided heads. The two bars to be compared are supported on a carriage running on two rails fixed between the pillars, so that first one bar and then the other can be brought underneath the microscopes. In order to keep the bars at a constant temperature, they are immersed in a water bath.

When using this instrument one bar is brought beneath the microscopes, and the cross wires are adjusted till they exactly coincide with the image of the division lines on the bar. The other bar is then substituted, and the number of turns and fractions of a turn of the micrometer screws necessary to bring the cross wires into coincidence with the image of the division lines is noted. Preliminary experiments are made to determine the magnification of the microscopes and the pitch of the micrometer screws, so that from the number of revolutions the difference in length of the two bars can be calculated. In the instrument in use for comparing the standard metres at the Bureau International des Poids et Mesures at Paris, one division on the micrometer heads corresponds to a difference in length of the bars of 0.001 millimetre, *i.e.* to one-millionth of a metre.

20. The Cathetometer.—In order to measure a vertical height, an operation of frequent occurrence in Physical measurements, an instrument called a cathetometer is usually employed. One form of cathetometer is shown in Fig. 7. A vertical metal pillar, AB, is fixed in a heavy tripod-stand in such a manner that it can be rotated about a vertical axis. This pillar has a divided scale engraved along one face. Two carriages C and D slide along the pillar. One of these, C, carries a telescope, T, while the other, D, has a clamping screw, by means of which it can be clamped to the pillar in any position. These two carriages are connected together by a fine screw, E, so that, D being clamped to the pillar, by turning this screw the other carriage, C, together with the telescope, can be moved through a small distance and its position accurately adjusted. The position of the carriage C is read off on the scale by means of a vernier, V. A spirit-level, L, serves to show when the axis of the telescope is horizontal, a screw, F, being used to make this adjustment.

Fig. 7.

When using the instrument to measure the vertical distance between two points, the pillar is first set vertical by means of the levelling screws, this adjustment being complete when on rotating the pillar the position of the bubble of the spirit-level L does not alter. The carriages are then moved till the lower end of the object to be measured is seen through the telescope. The carriage D is then clamped, and by turning the screw E the intersection of the cross wires of the telescope is made to coincide with the image of the lower point. The position of the carriage C having been read by means of the vernier, the carriage is moved till the image of the upper point coincides with the intersection of the cross wires. The

difference between the two readings gives the vertical distance between the two points.

In order to obtain a correct result it is very important that the axis of the telescope in the two positions should be exactly parallel. This will be evident from Fig. 8, where the axis of the telescope when at T_2 is shown inclined to its position when at T_1. The distance between X and Y would then be read off as T_1T_2 instead of T_1C as it ought to be. The error due to this cause is minimised by always setting the axis of the telescope horizontal, as shown by the delicate spirit-level L, by means of the screw F before making the final adjustment of the slide in *both* positions of the telescope.

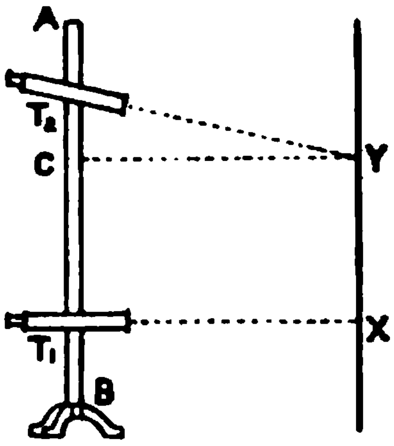


FIG. 8.

21. Units of Surface.—For all scientific purposes the unit of surface is a square of which each side is of unit length. In the *c.g.s.* system, therefore, the unit of surface is a square, each side of which is one centimetre, and is called a square centimetre. A square centimetre is sometimes written sq. cm. and sometimes cm². For measuring such surfaces as would in the British system be measured in acres, the unit employed in the metric system is the hectare, which is 10,000 square metres; one hectare is equal to 2.471 acres. The dimensions of surface are $[L^2]$.

22. Measurement of Surface.—The area of certain figures can be readily calculated by geometry. Some of the commonly occurring cases are given in the following table:—

Figure.	Area.	Remarks.
Square . . .	a^2	a = length of side.
Rectangle . .	$a.b$	a and b = length of two adjacent sides.
Parallelogram .	$a.b$	a = height, b = base.
Triangle . . .	$\frac{1}{2}a.b$	a = height, b = base.
Circle . .	$\pi.r^2$	r = radius, $\pi = \frac{\text{circumference}}{2r} = 3.1416$.
Ellipse . . .	$\pi.a.b$	a and b = semi-axes.
Sphere (surface of) .	$4\pi.r^2$	r = radius.

The surface of an irregular plane figure may be determined experimentally by tracing the outline of the figure on a sheet of cardboard or tinfoil, then cutting it out and weighing. The weight of a square centimetre of the same card or foil is then measured, and from this the area of the figure calculated from its weight. Another method in common use is to trace the figure on paper (called curve or squared paper) which is subdivided by two series of parallel lines at right angles to one another into a number of equal small squares. The number of these squares in-

cluded within the figure is then counted, and by multiplying this number by the area of each of the squares, the area of the figure is determined.

For an account of the rules for approximately calculating the area of certain figures, and for a description of instruments for mechanically obtaining the area of plane figures, reference must be made to text-books on mensuration and the integral calculus, since they cannot be profitably described without assuming a knowledge of the calculus.

23. Units of Volume.—The unit of volume for all scientific purposes is the volume of a cube each edge of which is of unit length. Thus in the *c.g.s.* system the unit is the volume of a cube each edge of which is one centimetre in length. This unit is called the cubic centimetre, and is generally written c.c. or cm^3.

For commercial purposes the unit of volume in the metric system is the litre, which is the volume of a kilogram of pure water at the temperature of its maximum density (4° C.). The litre is thus for all practical purposes equal to 1000 cubic centimetres or one cubic decimetre. One litre is equal to 1.76077 imperial pints, or 0.220097 gallon.

The following table is convenient for converting pints to litres, or *vice versa* :—

British to Metric.	Metric to British.
Pints = Litres.	Litres = Pints.
1 = 0.5679	1 = 1.7608
2 = 1.1359	2 = 3.5215
3 = 1.7038	3 = 5.2823
4 = 2.2717	4 = 7.0431
5 = 2.8396	5 = 8.8039
6 = 3.4076	6 = 10.5646
7 = 3.9755	7 = 12.3254
8 = 4.5435	8 = 14.0862
9 = 5.1114	9 = 15.8469

The dimensions of volume are $[L^3]$.

The following table gives the volumes of some of the simpler geometrical figures :—

Figure.	Volume.	Remarks.
Cube	a^3	a = length of edge.
Rectangular parallelopiped . . .	$a.b.c$	a, b, c = lengths of three adjacent edges.
Sphere	$\frac{4}{3}\pi r^3$	r = radius.
Cylinder or prism . .	$A.h$	A = area of base. h = height.

The experimental measurement of volume will be considered later (§ 146).

The discussion of the methods of measuring mass is for the present deferred (see § 95).

PART II—KINEMATICS

CHAPTER IV

POSITION

24. Province and Subdivisions of Mechanics.—The title mechanics is generally given to that part of physics which deals with the effects of force on matter, without in any way considering *how* the force originates. *For the present* we may regard force as typified by muscular exertion. When we exert our muscular powers to overcome some obstacle we derive, by means of our sense organs, a certain sensation which we describe as due to the fact that we are exerting a force. When any inanimate agency produces effects on bodies which are similar to those which we produce by muscular exertion, it is in the same way said to exert force.

As far as mechanics is concerned, the effects of force on matter are of two kinds—(1) change of motion, and (2) change of size or shape.

Before studying the effects of force on the motion of bodies, which constitutes the branch of mechanics called *Dynamics*, it is advantageous to study motion in the abstract, *i.e.* without reference to the cause of the motion. This branch of mechanics is called *Kinematics.*

25. Material Particle.—A portion of matter so small that, for the purposes of the discussion in hand, the distances between its different parts may be neglected, compared to the other lengths we are considering, is called a *material particle.*

The limiting size of a material particle varies very much in different investigations. Thus in some astronomical problems the earth and the other planets are treated as material particles, while if we attempt to account for the different kinds of light emitted by glowing gases, by a consideration of the vibrations of the molecules or even of the atoms, it is no longer permissible to regard an atom as a material particle.

26. Position.—The definition of a material particle amounts to a statement that the position of such a material particle can be represented by a geometrical point, which has position but not magnitude. This at once leads to the question of position.

In order to define the position of a point, we require to know its distance from some fixed point of reference, called the origin, and also the

direction in which we must go in order to pass from the origin to the given point. In order to be able to specify this direction, it is necessary that we have some fixed direction. Suppose we first take the case of the definition of the position of a point on a plane surface. Let P (Fig. 9) be such a point, and let O be the origin, and OX (called in geometry the initial line) be the fixed direction.

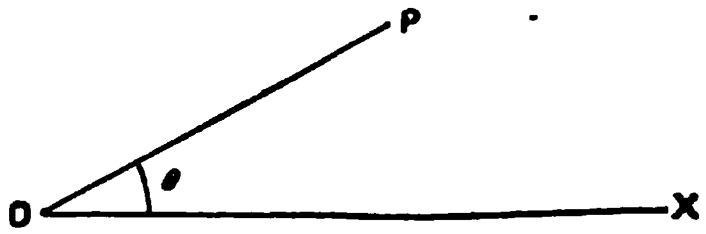


FIG. 9.

Then it is evident that if we know the angle θ, which the straight line joining P to the origin makes with OX, and also the distance (r) we have to travel along this line from O to reach P, then the position of P is completely defined. The quantities r and θ, which serve to define the position of P, are called the co-ordinates of P.

Another method of defining the position of a point in a plane is to have two fixed intersecting straight lines, called the axes, inclined at any angle to one another, and refer the position of the point to these lines. Thus, suppose we have two fixed straight lines, XOX′ and YOY′ (Fig. 10), intersecting at O (the origin), and through any given point P we draw two lines, PN PM, parallel respectively to the axes, then if we are given the distances $\overline{NP}$ and $\overline{MP}$, the position of P is defined.

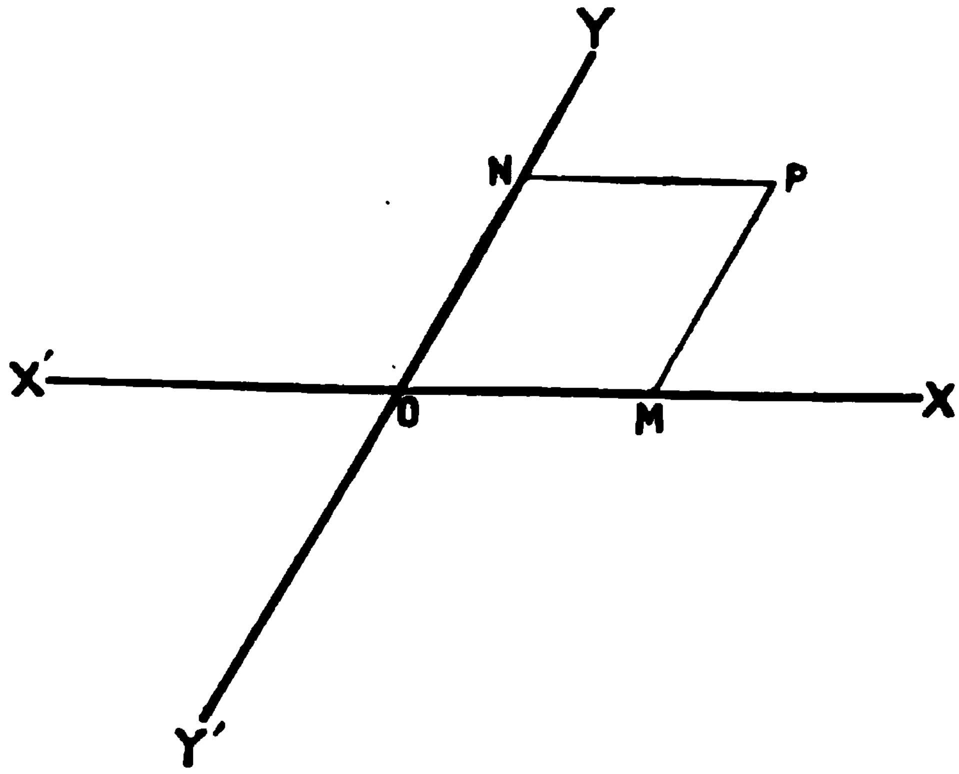


FIG. 10.

For if we measure off from O along OX a distance $\overline{OM}$ equal to $\overline{NP}$, and through N draw a line parallel to YY′, the point P must lie somewhere on this line. In the same way P must lie somewhere on the line NP, and hence must lie at the *only* point which is common to the two, that is at their point of intersection. It is usual to indicate the distance $\overline{OM}$ or $\overline{NP}$ by the symbol x, and $\overline{ON}$ or $\overline{MP}$ by the symbol y, so that the co-ordinates of the point P are x and y. In almost all practical applications of this method of defining the position of a point (called the Cartesian method) the two axes are taken at right angles to one another. In order to define the position of a point in *space* we require three co-ordinates. In the Cartesian method three axes are taken which are at right angles to each other, and the co-ordinates of a

point are then the distances from the origin of the feet of the perpendiculars drawn from the point to the three axes.

27. Vectors and Scalars.—Suppose we have the positions of two points (O and P) given. Then the position of P relative to O is given by the *length* and the *direction* of the straight line $\overline{OP}$ drawn from O to P. That is, starting from O you will arrive at P if you go in the direction of the line $\overline{OP}$ for a distance equal to the length of this line.

In geometry the expression $\overline{OP}$ is used simply to designate a line. When, however, it is used to designate the *operation* by which the line is drawn, *i.e.* the motion of a tracing point in a certain definite direction for a certain definite length, it forms an example of a quantity called a *vector*. To emphasise this fact we shall indicate a line such as $\overline{OP}$, when it is used as a vector, by an arrow placed over the letters which define the ends of the line, thus $\overrightarrow{OP}$. The arrow will here remind us of the distinctive property of a vector, namely, that in addition to a definite magnitude, it has also a definite direction, for we are constantly in the habit of indicating a direction by means of an arrow-head. The expressions $\overrightarrow{OP}$ and $\overrightarrow{PO}$ represent two different vectors, for although the distance is the same in the two cases, yet in one the tracing point is supposed to move from O to P, and in the other from P to O. Where we use a single symbol to represent a vector quantity, and we want to emphasise that it is a vector, we shall use a thick fount of type, while for scalar quantities the ordinary type will be employed. Thus $\mathbf{v}$ will represent a vector of which the magnitude is v units in some definite direction.

A quantity which has only magnitude and not direction is called a *scalar*. Thus mass and density are scalars, but velocity and force, as we shall see, are vectors, for they have not only magnitude, but have associated with this magnitude a certain direction.

28. Motion.—If the position of a material particle is changed, then if we only consider its state before and after the process of change, and take no account of the time during which this change takes place, we are said to study the displacement of the particle. When a particle is displaced, however, from one point to another, it must travel over a continuous path from one position to the other; and further, it must take a certain time in travelling over this path, so that it has occupied in succession every point along this path. When we consider the actual process of change of position as occurring during a certain time, we are said to study the *motion* of the particle, while that branch of mechanics which is concerned with the motion of bodies treated in the abstract, *i.e.* without considering what causes the motion or change of motion, is called *Kinematics*.

29. Different Kinds of Motion.—The motion of a material particle, taken with reference to some fixed point as origin, can consist either in change in the distance of the particle from the origin, the

direction of the straight line joining the particle to the origin remaining fixed, *i.e.* motion can take place along this straight line either away from or towards the origin, or in a change in the direction of the line joining the particle to the origin, the length of this line remaining fixed, *i.e.* motion along the circumference of a circle having the origin as centre, or in a combination of these two. In the case of a material particle, since it has no parts, the above are the only kinds of motion possible, and this form of motion is called *motion of translation*. If, however, instead of dealing with a material particle, we are dealing with a body of appreciable size, so that its different parts can have different motions, we have a further kind of motion possible. Thus in addition to a motion of translation, in which the body moves so that the line joining any two points in the body is always parallel to some fixed line, the body may spin or rotate. In the case of a pure translation, the motion of all the particles, of which we may consider the body to be built up, is exactly the same, while when the body rotates the motions of the different parts of the body are different. The most general kind of motion of which an extended body is capable is a combination of a rotation with a translation.

As an example of a motion of translation, if we neglect the curvature of the earth's surface, we may take the case of a boat sailing in a straight line. The fly-wheel of a stationary engine is an example of a motion of pure rotation. The motion of the screw propeller of a ship, the wheel of a locomotive, and a ball rolling along the ground are obvious examples of the combination of a motion of translation with one of rotation.

CHAPTER V

MOTION OF TRANSLATION

30. Velocity, Speed.—The rate[1] at which a point changes its position is called its velocity. From what has been said in § 27 it is evident that the change in the position of a particle must not only have magnitude, *i.e.* there must be a certain distance measured along the path traversed by the particle between its first and last positions, but also the motion of the particle must have been in some direction, although not necessarily along a straight line, so that velocity is a *vector*. Velocity, therefore, may vary both in regard to its magnitude and also in regard to its direction. This may be illustrated by the motion of a train going round a curve. Here, although the magnitude of the velocity may be constant, *i.e.* the train may travel along the rails for equal distances in each successive second, yet the direction of the motion is continually varying, since at any given point it is along the tangent to the curve at that point.

Hence, to measure the velocity of a particle two things have to be determined: (1) the space which the particle has moved over in a given time, and (2) the change in the direction of motion during this time. In ordinary language, and in very many books on mechanics, the word velocity is used to indicate the first of these rates, *i.e.* the *space* passed over in a given time, without taking any account of any change in direction which may take place. Thus the end of the hand of a watch is said to move with uniform (*i.e.* constant) velocity, since it moves over equal spaces in successive equal times. It is, however, evident that the direction of the velocity is continually altering, and hence from this point of view the velocity is variable. It therefore saves confusion if we use, at any rate wherever ambiguity may arise, a separate word to denote the rate at which a particle describes its path, without reference to the direction, and for this purpose the word *speed* is generally used. Hence, if a particle moves in a straight line (so that the direction of motion remains constant), and passes over equal spaces in successive equal times, its *velocity* is said to be constant. If, however, a particle moves in a curve, so that its direction of motion continually changes, but passes over equal

[1] The ratio of the total change in any quantity which occurs during a given time to that time is called the *rate* at which that quantity is changing.

lengths of its path in successive equal times, its *speed* is said to be constant or uniform.

Constant speed is measured by dividing the space passed over in any given time by that time. Thus, suppose a particle passes over a space, s, in a time, t, then the speed, v, is given by the equation

$$v = s/t.$$

Hence unit speed is such that unit space is passed over in unit time. In the *c.g.s.* system unit speed is such that one centimetre is passed over in one second. The unit of speed has not received any recognised name, but when a particle passes over say 10 cm. in every second it is said to have a speed of 10 centimetres per second. This is often written 10 cm./sec. The only speed which has a recognised name is that of one nautical mile per hour, which is called a knot.

The dimensions of speed can be obtained from the equation

$$v = s/t$$

by writing in the symbols for the units, when we get

$$v[V] = s[L] / t[T].$$

If s and t are each unity, then v is also unity, and we get the dimensional equation

$$[V] = [LT^{-1}],$$

from which it follows that the dimensions of a speed are $[LT^{-1}]$.

The speed, v, is considered to be positive if the particle is moving *away* from the origin from which the distances are measured, and negative if the particle is moving towards the origin. Thus, if we measure vertical distances from the surface of the earth, the speed of a balloon is positive when it is ascending and negative when it is descending, while a bucket being lowered down a well has positive speed, but when being raised it has negative speed.

31. Variable Speed.—If a particle moves over unequal spaces in successive equal intervals of time, its speed is variable. Variable speed, at any instant, is measured by dividing the space passed over in a time, including the given instant, so small that during this interval the speed does not appreciably alter, by this interval of time. Suppose that during the very small interval of time which we may indicate by δt[1] the particle

[1] We here use δt as a convenient symbol for a very small interval of time, or in other words, the symbol δ is used to indicate a small increment in the quantity to which it is prefixed. The expression δt must not be looked upon as the product of δ and t, but as a single expression, in fact a kind of shorthand expressing a *small* increment of time. In the same way δs represents a small increment in the length, measured from some fixed point, of the path traversed by the particle. Suppose, then, that during the time t the particle has passed over a distance s, and at the time $t + \delta t$ it has passed over the space $s + \delta s$, it is evident that it has passed over the space δs in the time δt, and its speed is $\dfrac{\delta s}{\delta t}$.

moves over the space δs, then the speed with which the body is moving at the time when the observation is made is given by

$$v=\frac{\delta s}{\delta t}$$

It is important to be quite clear as to exactly what the above equation implies. In the first place it follows, from the definition of uniform speed given in § 30, that if a particle passes over a distance s in a time t, then the speed is given by s/t, whatever the value of t may be. Thus if a particle moving with a uniform speed pass over a metre in a second, the speed is one metre per second. It will pass over a tenth of a metre in a tenth of a second; the speed, however, will be the same (*i.e.* $\frac{1}{10}/\frac{1}{10}=$ 1 m/sec.). In the same way it will pass over a millionth of a metre in a millionth of a second, and the speed will be as before, 1 metre per sec. $\left(\frac{1}{10^6}\bigg/\frac{1}{10^6}\right)$. Hence if we were able to measure or calculate the space passed over in this small interval of time, we should obtain the same value for the speed as would be obtained if we measured the space passed over in one second or a thousand seconds. Thus the value obtained for the variable speed ($v=\delta s/\delta t$) is such that if we had a particle moving with a *constant* speed, such that it passed over the space δs in the time δt, it would in one second traverse a space equal to v units of length. We might therefore say that the variable speed of a particle is measured at any instant by the space passed over in one second by a second particle moving uniformly with a speed equal to that with which the first particle is moving at the given instant.

A consideration of these two definitions will assist in making the matter clear; any difficulty which may be encountered may be lessened by recollecting that every one probably has some idea what they mean by saying that at any instant a train is travelling at, say, fifty miles per hour, though probably the train may not actually travel more than a mile or so in all.

32. Acceleration.—The word acceleration is, in its most general sense, used to indicate any change in velocity. Hence it may mean an increase or a decrease in the speed, or a change in the direction of motion. In this sense acceleration is clearly a vector. It is, however, common to use the term acceleration with reference to the change in speed only, when it is a scalar.

Acceleration may be uniform or variable. In uniform acceleration equal changes of speed occur in equal times.

If the speed is increasing, then the acceleration is positive, while if the speed is decreasing, the acceleration is negative. Hence negative acceleration is what in ordinary language is called a retardation.

Uniform acceleration is measured by the change in speed that takes place in a given time divided by that time. Hence if the speed of a

particle change during the time t from v_1 to v_2, the acceleration (a) is given by the equation

$$a=\frac{v_2-v_1}{t}.$$

If the change in speed is unity, and takes place in unit time, we have unit acceleration. Thus if the speed of a particle increase in one second by one centimetre per second, its motion has unit acceleration. This is the unit of acceleration in the *c.g.s.* system. If the change of speed in one second is x centimetres per second, then the acceleration is x centimetres per second per second,[1] which may be written x cm/sec.^{2}.

Suppose the *change* in speed in a time t to be v, then the acceleration is given by

$$a=v/t$$

If in this equation we introduce symbols for the units, we get

$$a[A]=v[V]\div t[T].$$

Hence making a, v, and t each unity, we get the dimensional equation

$$[A]=[V]\div[T].$$

Substituting the dimensions of $[V]$ from § 30, we get

$$\begin{aligned}[A]&=[L^1T^{-1}]\div[T]\\&=[L^1T^{-2}].\end{aligned}$$

The dimensions of acceleration are therefore 1 as regards length and -2 as regards time, and we are reminded of this double appearance (*i.e.* to the second power) of the unit of time by the expression centimetre per *second* per *second*.

As in the case of variable velocity, variable acceleration is measured by dividing the change in velocity occurring in a time so small that the acceleration does not appreciably change during the interval by that time.

33. Velocity Curve.—Take two axes at right angles, OX, OY (Fig. 11), and divide OX into a number of equal parts, OM_1, M_1M_2, M_2M_3, &c., and suppose each of these parts to represent an equal interval of time, say one second, so that we are measuring time along the axis of X. Next suppose we have a particle which starts from rest, and at the points O, M_1, M_2, &c., we draw perpendiculars M_1P_1, M_2P_2, &c., to represent the speed of the particle at the instants of time represented by these points, *i.e.* at the commencement of the first, second, third, &c., second of motion. Since the particle starts from rest, the perpendicular at O is zero. If now the interval of time taken (*i.e.* a second in the above example) is sufficiently small so that the velocity has not greatly changed from one point to the

[1] An acceleration is sometimes referred to as of so many centimetres per second, or feet per second. This is wrong, since acceleration is a change in speed and not in position, as these expressions would lead us to suppose.

next, the straight lines joining the points O, P_1, P_2, P_3, &c., will form a continuous curve; and this curve is called the velocity[1] curve.

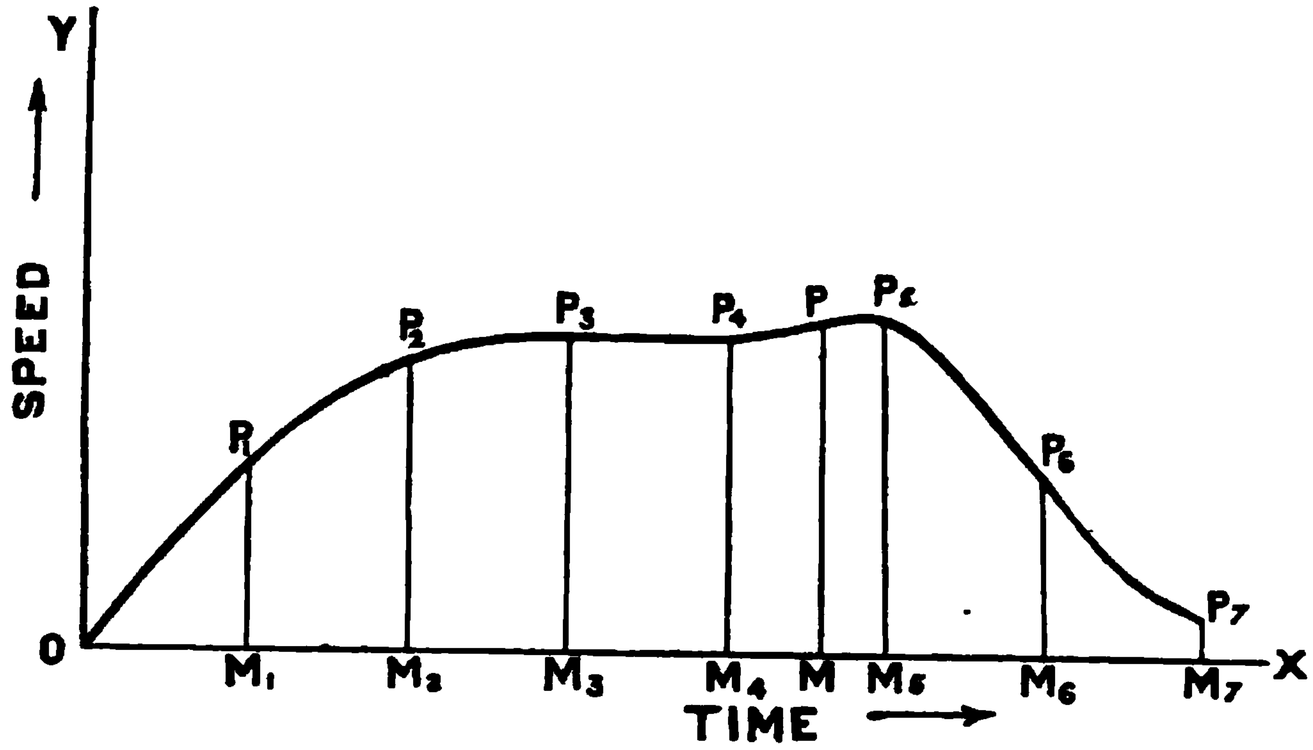


FIG. 11.

Having drawn this curve, if through any point M, corresponding to a time t, we draw a perpendicular to meet the curve at P, then MP represents the speed of the body at a time t from the start.

If the speed of the particle is uniformly accelerated, the speed will increase by an equal amount in each unit of time. Thus, suppose that the particle starts from rest and moves with an acceleration a. Let OM′ (Fig. 12) represent one second, then, since the speed of the particle at the end of a second is a, the ordinate $\overline{M'p'}$ representing the velocity at one second from the start is equal to a. If OM represents a time t, then, as the speed increases by an amount a in each second, the speed represented by $\overline{MP}$ will be at. If we join Op' and OP, since

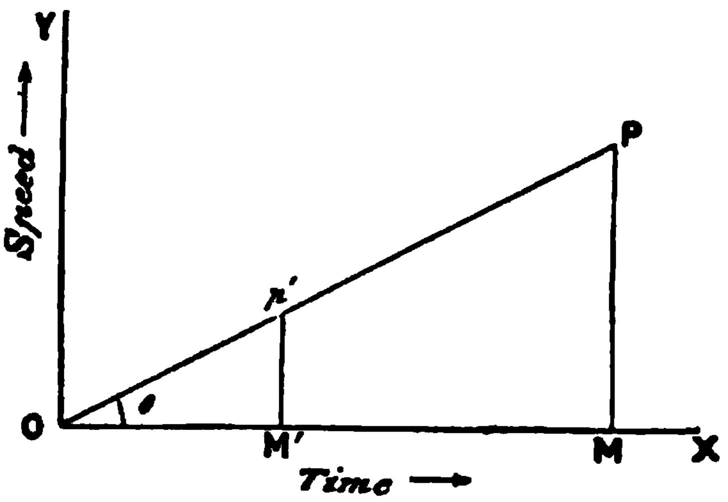


FIG. 12.

$$\frac{\overline{p'M'}}{\overline{OM'}} = a = \frac{\overline{PM}}{\overline{OM}}$$

and the angles p'M′O and PMO are both right angles, it follows that the triangles p'OM′ and POM are similar, and the angles p'OM′ and POM are equal. The point p' must therefore lie on the straight line OP. In the same way it may be shown that all the extremities of

[1] Strictly it ought to be called the *speed* curve.

the ordinates which represent the speed of the particle at different instants must lie on the straight line OP, so that this line must be the velocity curve for the uniformly accelerated particle.

Since $\frac{\overline{PM}}{\overline{OM}}$ is the tangent of the angle, θ, which the velocity curve makes with the axis of time, it follows that the tangent of this angle is numerically equal to the acceleration with which the particle moves. If instead of starting from rest the particle start with an initial speed v_0 and move with an acceleration a, then the velocity curve will be represented by the straight line PP′ (Fig. 16). In this case OP is equal to v_0 and NP′ is equal to $a\ t$, if OM represents the time t. Hence the speed at a time t is given by

$$v = v_0 + at$$

for

$$v = \overline{MP'} = \overline{MN} + \overline{NP'} = \overline{OP} + \overline{NP'}.$$

34. Graphical Representation of the Space passed over by a Moving Particle.—The velocity curve for a particle moving with uniform speed is a straight line parallel to the axis OX, for in this case the lines P_1M_1, P_2M_2, &c., will all be of equal length. Let P P_1 P_2 (Fig. 13) be the velocity curve in the case of a particle moving with a uniform speed v. Then OP, N_1P_1, N_2P_2, &c., are each of such a length that they contain as many units of length, on some convenient scale, as there are units of speed in v.

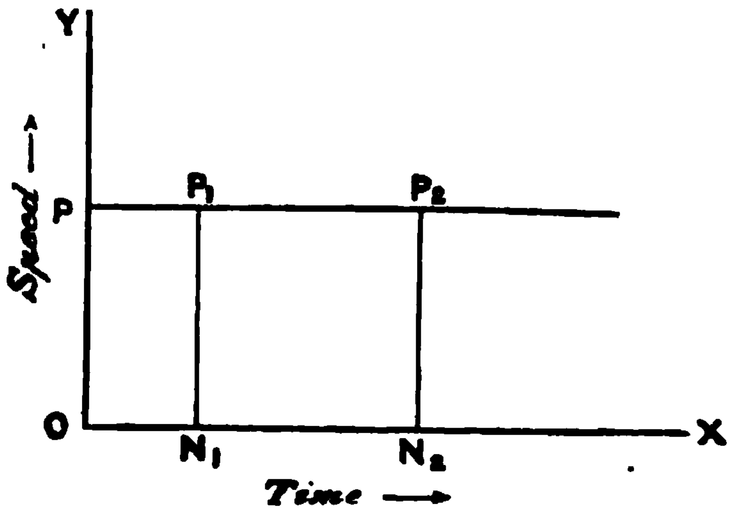


FIG. 13.

To find a graphical representation for the space passed over in a time t, take the point N_1 to correspond to the time T and the point N_2 to the time $T+t$, so that the length $\overline{N_1N_2}$ measured along the axis of time (OX) corresponds to the interval of time t. Now the area of the rectangle $N_1P_1P_2N_2$ is given by the product of the sides N_1P_1 and N_1N_2.

Therefore

$$\text{area of } N_1P_1\ P_2N_2 = \overline{N_1P_1} \times \overline{N_1N_2}$$

But

$$N_1P_1 = v, \text{ and } \overline{N_1N_2} = t$$

Hence

$$\text{area of } N_1P_1\ P_2N_2 = vt \quad . \quad . \quad . \quad \text{(i)}$$

But the space passed over in a time t is vt. Hence the space passed over by the particle during any interval may be represented by the area

of the rectangle contained between the ordinates drawn through the points corresponding to the commencement and end of the interval, the velocity curve, and the axis of time.

It is important to clearly understand the limitations under which the equation (i) is true. It will be at once seen, if we consider the dimensions of the two sides, that it is not a physical equation such as those we have been dealing with up to now, for the dimensions of one side are $[L^2]$ and of the other $[L^1T^{-1}.\ T^1]$ or $[L^1]$. What we mean by this equation is that if one side of the rectangle $N_1P_1P_2N_2$ has as many units of length in it as there are units of speed in the speed of the particle, and an adjacent side is of as many units of length as there are units of time in the interval considered, then there will be as many units of area in the rectangle as there are units of length in the space passed over in the given interval.

In the cases where the speed is variable the velocity curve is no longer parallel to the axis of X. If the speed, instead of varying continuously, is supposed to remain constant during each second at the value it has at the commencement of the second, but to change suddenly at the end of the second to the value it really has at the commencement of the next, the velocity curve would consist of a stepped line O M_1 P_1 Q_2 P_2 Q_3 P_3, &c. (Fig. 14). Since the speed during each second is uniform, we may apply the result just obtained for uniform speed to each period of a second separately, and the space passed over in any given time will be represented by the area included between this stepped line, the axis of time (X), and the two ordinates drawn through the points corresponding to the commencement and end of the interval. Since throughout each second the assumed speed is less than the actual speed, the space passed over, as represented by the above area, will be less than the true space. Next, suppose that the particle moves uniformly during each second with the speed it really has at the *end* of that second. In this case the space passed over will be represented by the area enclosed between the stepped line R P_1 R_1 P_2 R_2 P_3, &c., the axis of time, and the ordinates at the commencement and end of the time interval, and will be greater than the real space. The difference between the space passed over in these two cases, where the changes in speed are supposed to occur only at the end of each second, is the sum of the small rectangles which are shaded in the figure. Next,

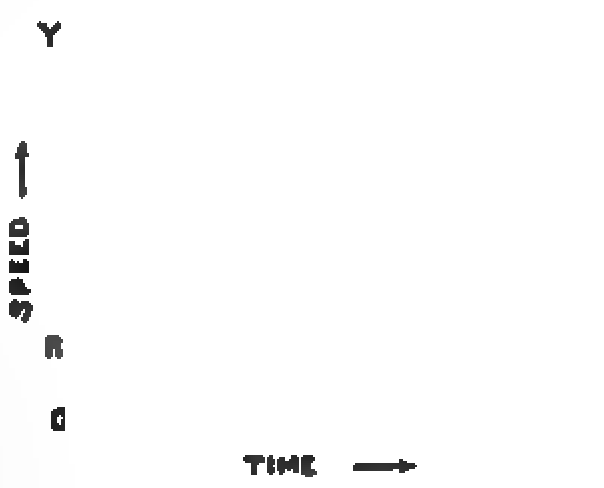


Fig. 14.

6

let us suppose that we take half-second steps, *i.e.* that the velocity remains constant for half a second and then changes abruptly. The velocity curves now obtained are shown in Fig. 15. As before, the difference between the space passed over when the particle is supposed to move during each half-second with the velocity it actually possesses at the end of the half-second, exceeds the space passed over when the particle moves with the velocity it actually possesses at the commencement by the area of the shaded rectangles. By comparing the area of these rectangles with those in Fig. 14, where whole second intervals were used, to facilitate which comparison the rectangles for the whole second periods have been dotted in on Fig. 15, it will be seen that the difference in area, and hence the difference in the space passed over, is much less in the second case than in the first. Also, the difference between the area enclosed by the two stepped lines respectively, the axis of times and any two ordinates and that included between the actual velocity curve, the axis of time and the same ordinates is much less than before. By taking quarter-second periods, and then tenths and hundredths of a second, we should find that the difference between the area corresponding to the case where the particle moves during each time interval with a constant velocity corresponding to the actual velocity it possesses at the commencement of the interval, and that corresponding to the case where the constant velocity during each interval corresponds to the actual velocity at the end of the interval, gets less and less. Hence we conclude that if we take an infinite number of intervals the two areas will be equal, and will coincide with that enclosed by the velocity curve O P_1 P_2 P_3, &c. In the limit, therefore, since the space actually passed over must be intermediate in value between that passed over in the two hypothetical cases, it follows that when these are equal the actual space passed over must also be equal to either, and must be represented by the area enclosed between the velocity curve O P_1 P_2 P_3, the axis of time and any two ordinates.

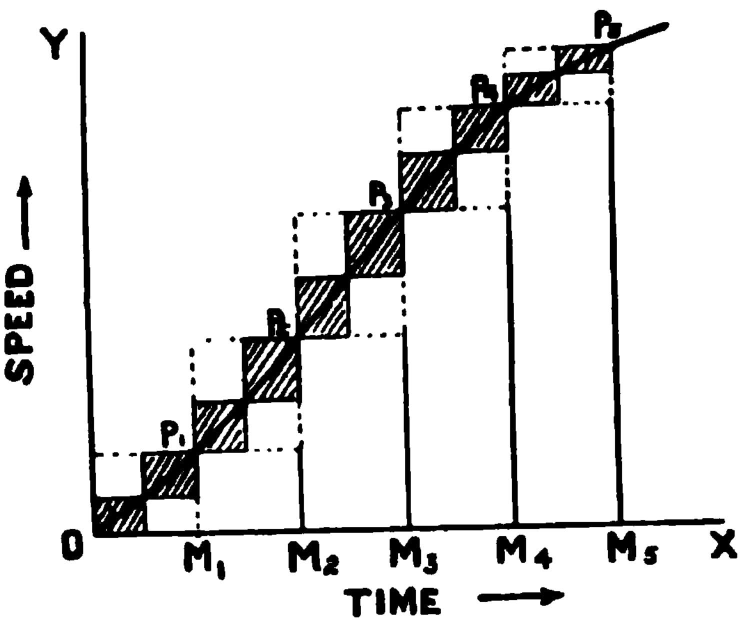


FIG. 15.

35. Space Passed over by a Particle when its Motion is Uniformly Accelerated.—Suppose that a particle starting with a speed v_0 moves with a uniform acceleration a for a time t.

In this case PP′ (Fig. 16) is the velocity curve, OP presenting the initial speed v_0, and P′M the final speed v_0+at. The space passed over in the interval t is represented by the area of the figure POMP′. Now

the area of the figure POMP′ is equal to the area of the rectangle POMN, together with that of the triangle PNP′. But the area of the rectangle POMN is equal to $\overline{PO} \times \overline{OM}$ or to $v_0 t$. The area of the triangle PNP′ is equal to $\frac{1}{2}$ $\overline{PN}.\overline{P'N}$ or $\frac{1}{2}\, t \times at$, since $\overline{P'N} = at$. Hence the total space passed over will be given by

$$s = v_0 t + \tfrac{1}{2}\, at^2.$$

If the particle start from rest v_0 is zero, and the space passed over in a time t is

$$= \tfrac{1}{2}\, at^2,$$

which is also evident from the figure, since the space passed over would then be represented by the triangle PP′N.

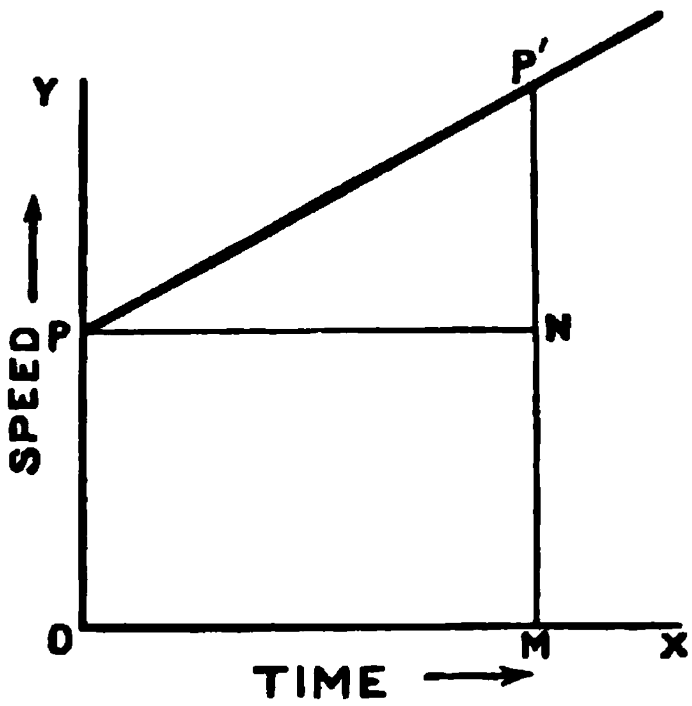


FIG. 16.

The initial speed being v_0, and the final speed $v_0 + at$, the mean or average speed is $\frac{1}{2}(2v_0 + at)$. If a particle were to travel for a time t with a uniform speed equal to this mean speed, the space passed over would be

$$s = \tfrac{1}{2}(2v_0 + at)t$$
$$= v_0 t + \tfrac{1}{2}at^2.$$

Thus in the case of uniformly accelerated motion the space passed over may be obtained by supposing that the body travels during the whole time with a uniform speed equal to the mean of the initial and final speeds. Hence if v_0 is the initial speed and v_t the speed after a time t, the space passed over during this time is given by

$$s = \tfrac{1}{2}(v_t + v_0)t.$$

If the body starts from rest

$$s = v_t t/2,$$

or

$$t = 2s/v_t.$$

Substituting this value for t in the expression $s = at^2/2$, we get

$$s = 2as^2/v_t^2,$$

or

$$a = v_t^2/2s,$$

an expression which we shall sometimes find useful.

36. Graphical Representation of a Velocity.—Up to the present we have been considering the speed of a particle exclusively; we have now to take into account not only the speed but also the direction of motion. In discussing all questions involving velocities, it is convenient

to represent velocities by straight lines. For this purpose we draw a straight line in the direction of the velocity, and of such a length that it contains as many units of length as there are units of speed in the velocity considered. In order to show in which sense the motion takes place it is usual to place an arrow-head somewhere along the line. Thus a velocity of 3.6 centimetres per second in a north-westerly direction would, if we take a centimetre to represent a velocity of one centimetre per second, and if SN (Fig. 17) represent the south to north direction, be represented by either of the lines AB or CD, which are each 3.6 cm. long and point in the proper direction. Similarly, any other line equal and parallel to AB will represent the given velocity. The reason why an infinite number of lines can be drawn to represent any given velocity, is that although velocity has magnitude and direction it has not position, hence we can take any point from which to draw a straight line to represent a given velocity.

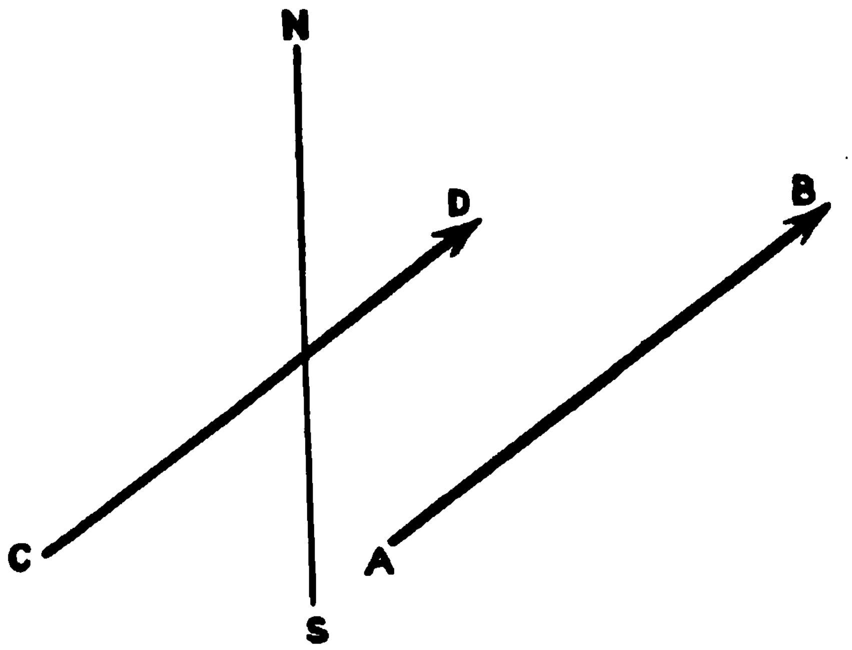


FIG. 17.

37. Composition of Velocities.—We have hitherto supposed the axes to which we have referred the velocity of a particle to be fixed in space, and our results will not be affected although it is impossible to realise such fixed axes, since it is immaterial, as far as the relative movement of any particle with respect to certain fixed lines or points is concerned, whether these are themselves in motion or at rest, and it is with relative motion that we are exclusively concerned. We now pass to the consideration of the case where the axes to which we refer the motion of a particle are themselves in motion with reference to some other axes, which we take for the nonce to be fixed.

As an example we may take the case of a man walking along the deck of a ship, which is moving in a given direction with reference to the surface of the sea. We might measure the velocity with which the man is moving with reference to the ship, *i.e.* say he was walking at the rate of 200 centimetres per second in a direction inclined at 45° to the length of the boat and towards the bow. The ship itself (and therefore the axes we have used to define the man's velocity) is, let us suppose, travelling at the rate of 500 centimetres per second due north, and that we require to find the velocity with which the man is travelling with reference to the surface of the sea.

If the man remained standing still on the deck while the ship moved for one second, he would be displaced 500 centimetres to the north, *i.e.* from O to N (Fig. 18). This represents in the given case the displacement of the co-ordinates by which the man's position with respect to the ship is defined. If now the ship were to remain stationary while the man moved for one second with his given velocity, *i.e.* 200 centimetres in the north-east direction, he would arrive at R. If the two movements were to go on simultaneously, he would therefore during one second travel from O to R. His actual path with reference to the surface of the sea is along OR, and he will have moved over this distance in one second. Now the man would have traversed the same path in the same time if, instead of moving with a velocity of 500 centimetres per second towards the north, represented by $\overrightarrow{ON}$ (Fig. 17), and at the same time with a velocity of 200 centimetres towards the north-east, represented by $\overrightarrow{OP}$, which is equal and parallel to $\overrightarrow{NR}$, he had travelled for one second with a single velocity represented by $\overrightarrow{OR}$. This single velocity $\overrightarrow{OR}$, which would give the same displacement as the two given velocities, is called their resultant, and they are called the component velocities. Hence we have the rule for finding the resultant of two velocities :—From any given point (O) draw two straight lines to represent the given velocities in magnitude and direction,[1] and complete the parallelogram with these lines as adjacent sides. Then the diagonal of this parallelogram drawn through O will represent the magnitude and direction of the resultant velocity. This result is called the Parallelogram of Velocities.

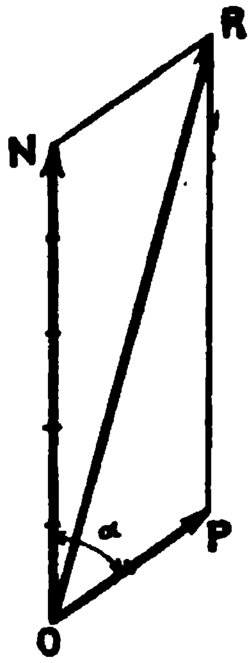


Fig. 18.

What we are really doing in the parallelogram of velocities is to *add* together two velocities, and we shall find that whenever we add two quantities which are *vectors* we shall make use of the parallelogram law.

The parallelogram of velocities enables us to find the resultant of any number of velocities, since by its means we can replace any two of them by their resultant. Next, this resultant can be combined with one of the remaining velocities, and so on till a single resultant is left.

A consideration of Fig 18 will show that the resultant velocity may be obtained without actually drawing the parallelogram ONRP, for if from the point N at the extremity of the straight line ($\overrightarrow{ON}$), which represents in

[1] The sense of the two velocities must be either both towards O or both away from O, *i.e.* *both* velocities must be represented as directed, away from O or towards O. In the first case the sense of the resultant will be away from O, and in the second case towards O.

magnitude and direction one of the component velocities, we draw a straight line ($\overrightarrow{\text{NR}}$) to represent in magnitude and direction the other component velocity, then the straight line ($\overrightarrow{\text{OR}}$) joining the point O to the point R will represent in magnitude and direction the resultant velocity. Of course, the same result would be obtained by drawing through P, the extremity of the line representing the second velocity in the above case, a line $\overrightarrow{\text{PR}}$ to represent the first. This method of effecting the composition of velocities is generally referred to as the triangle of velocities, and leads to a convenient method of finding the resultant of a number of velocities, say, v_1, v_2, v_3, v_4. Suppose the given velocities are represented by the lines $\overrightarrow{\text{AB}}, \overrightarrow{\text{CD}}, \overrightarrow{\text{EF}}, \overrightarrow{\text{GH}}$ (Fig. 19), then from any point O draw a straight line $\overrightarrow{\text{OP}}$ equal and parallel to $\overrightarrow{\text{AB}}$, from P draw $\overrightarrow{\text{PQ}}$ equal and parallel to $\overrightarrow{\text{CD}}$, from Q

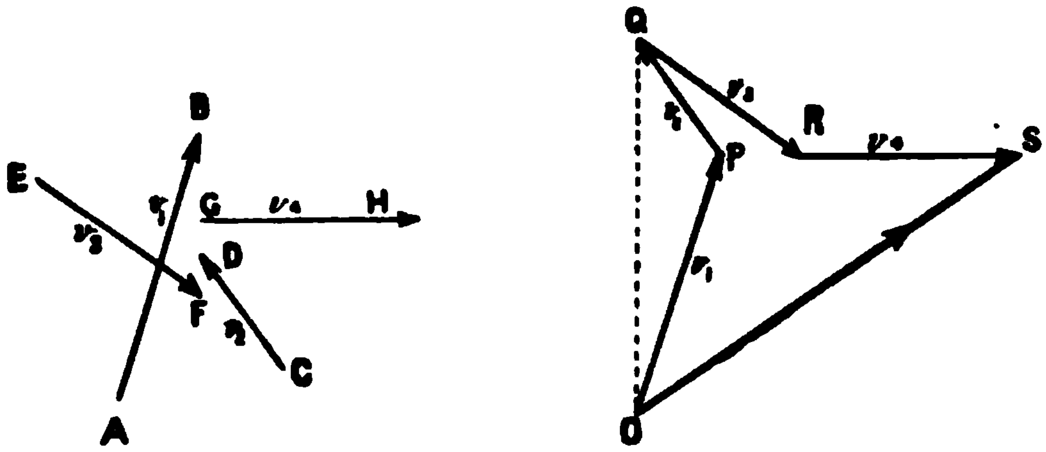


FIG. 19.

draw $\overrightarrow{\text{QR}}$ equal and parallel to $\overrightarrow{\text{EF}}$, and from R draw $\overrightarrow{\text{RS}}$ equal and parallel to $\overrightarrow{\text{GH}}$. Then the straight line $\overrightarrow{\text{OS}}$ will represent the resultant of the four velocities v_1, v_2, v_3, v_4. The velocity v_1 is represented by $\overrightarrow{\text{OP}}$ just as much as by $\overrightarrow{\text{AB}}$; the same remark applies to the other velocities. Hence by the triangle of velocities the resultant of v_1 and v_2 is represented by $\overrightarrow{\text{OQ}}$. Also the resultant of $\overrightarrow{\text{OQ}}$ and v_3, *i.e.* $\overrightarrow{\text{QR}}$, is represented by $\overrightarrow{\text{OR}}$, and so on. It is important to note that the value of the resultant obtained is independent of the order in which we draw the lines representing the velocities. Thus in Fig. 20 the velocities have been combined in the order v_4, v_1, v_3, v_2; the resultant, however, is the same as before.

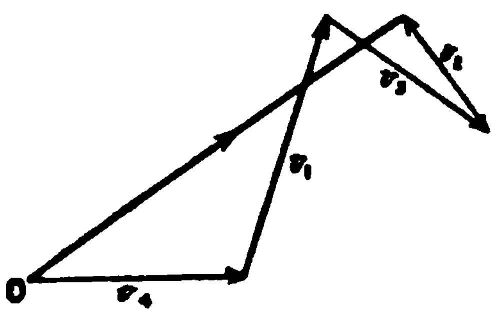


FIG. 20.

38. Resolution of Velocities.—It is sometimes convenient to replace the actual velocity of a body by two or more other velocities, which are so chosen that the actual velocity is the resultant of all these assumed

velocities. In such a case we are said to resolve the given velocity into component velocities. Suppose we require to resolve the velocity represented by $\overrightarrow{OP}$ (Fig. 21) into two component velocities, one directed along OX and the other along OY. If through P we draw PN parallel to OY, and PM parallel to OX, then we know by the parallelogram of velocities that two velocities, represented by $\overrightarrow{OM}$ and $\overrightarrow{ON}$ would have a resultant $\overrightarrow{OP}$, so that we may replace the velocity $\overrightarrow{OP}$ by the two velocities $\overrightarrow{OM}$ and $\overrightarrow{ON}$. Any given velocity can in this way be resolved in any number of separate ways; thus in Fig. 22 $\overrightarrow{OP}$ is resolved into the two velocities $\overrightarrow{OM}$ and $\overrightarrow{ON}$.

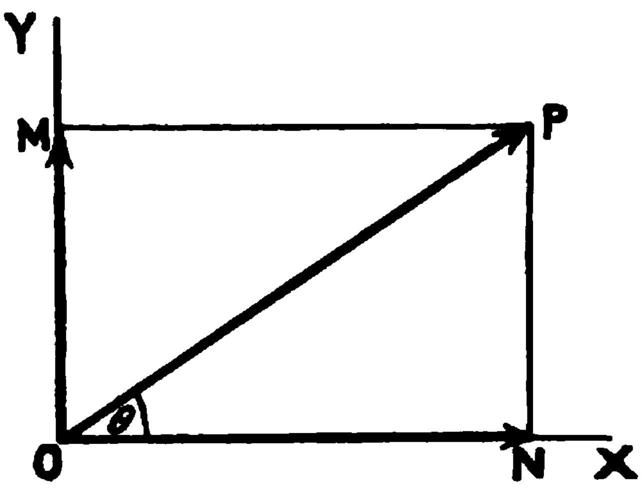


FIG. 21.

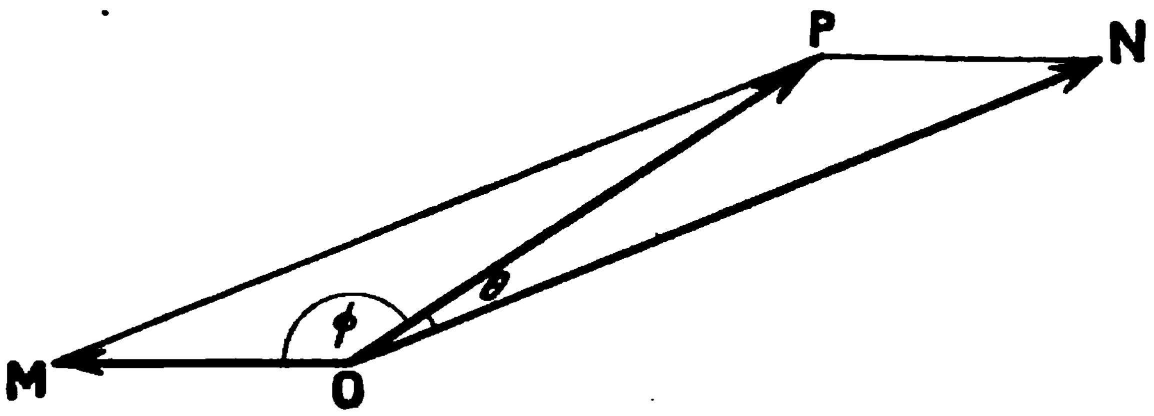


FIG. 22.

In the case where the components are at right angles to one another, if we call the angle POX (Fig. 21), between the direction of the resultant velocity and one of the components $\overrightarrow{ON}$, θ we have—

$$\frac{\overline{ON}}{\overline{OP}}=\cos\theta,$$

and
$$\frac{\overline{OM}}{\overline{OP}}=\cos(MOP)=\cos(90^\circ-\theta)=\sin\theta.$$

Hence if R is the velocity along $\overrightarrow{OP}$, and P and Q are the components of R along OX and OY respectively,

$$P=R\cos\theta,$$

and
$$Q=R\sin\theta.$$

When (as in Fig. 22) the two components are not at right angles, then if P is the component along $\overrightarrow{ON}$, making an angle θ with R, and Q that

along $\overrightarrow{OM}$, making an angle ϕ with R, we have, by a well-known proposition in elementary trigonometry—

$$
\begin{aligned}
P^2 &= R^2+Q^2-2PQ\cos\phi,\\
Q^2 &= R^2+P^2-2PQ\cos\theta,\\
R^2 &= P^2+Q^2-2PQ\cos OMP,\\
&= P^2+Q^2+2PQ\cos(\theta+\phi),
\end{aligned}
$$

where $\theta+\phi$ is the angle included between the components P and Q.

In practice, however, it is generally convenient to resolve a velocity into a component in the direction in which for the time being we may be confining our attention, and a second component at right angles to the first, which will obviously have no influence on the motion in the direction considered. Suppose, for instance, we require to find the time a sailing ship will take to go a certain distance in the teeth of the wind, being given that it sails at an angle of 45° to the direction of the wind with a speed of ten knots.[1] Let $\overrightarrow{NO}$ (Fig. 23) represent the direction of the wind, and $\overrightarrow{OP}$ the direction and magnitude of the velocity with which the boat sails. Resolve $\overrightarrow{OP}$ along ON, and at right angles to ON, *i.e.* along OE. Then we may consider that the boat moves with a velocity $\overrightarrow{OM}$ in the required direction, and with a velocity $\overrightarrow{OL}$ in a direction at right angles, which has no effect on the space passed over in the required direction. As a matter of fact the boat would sail alternately in the direction OP and in the direction OQ. The resolved part of the velocity along ON would be the same in the two cases, but the resolved part at right angles to ON would be alternately in the direction $\overrightarrow{OE}$, and in the direction $\overrightarrow{OW}$; and hence the space passed over at right angles to ON, due to each of these components, would be in opposite directions, and would neutralise each other. By geometry we see that, since PMO is a right-angled isosceles triangle, $\overline{OM}=\overline{OP}/\sqrt{2}$. Hence the component of the velocity of the boat in the direction ON is $10/\sqrt{2}$ knots. The time taken to cover x nautical miles in the direction ON will therefore be $x\div 10/\sqrt{2}$ or $\frac{\sqrt{2}.x}{10}$ hours.

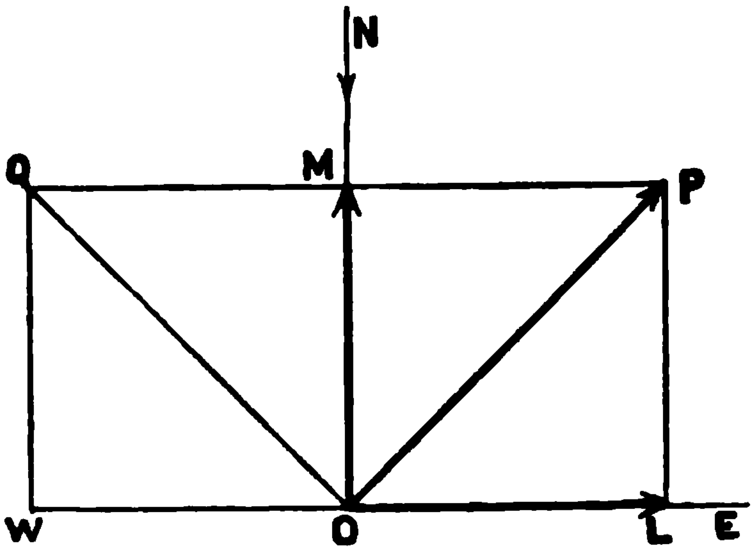


FIG. 23.

[1] The knot is the only special name we have for a unit of speed, and it represents a speed of one nautical mile per hour.

39. Composition and Resolution of Accelerations.—A uniform acceleration, being a vector or directed quantity, may be compounded or resolved in exactly the same way as a uniform velocity. Hence we have the parallelogram, triangle, and polygon of accelerations. It is, however, unnecessary to discuss these separately, since all we have said with reference to velocity applies (*mutatis mutandis*) to acceleration.

40. Composition of a Uniform Motion with a Uniformly Accelerated Motion.—In the previous sections we have dealt with the composition of two uniform velocities or of two uniform accelerations; it is, however, possible to compound a uniform velocity in one direction with a uniform acceleration in another. Thus in the example considered in § 37 we might suppose that while the man walked with uniform velocity with reference to the deck, the ship itself was moving with an accelerated motion with reference to the surface of the sea.

As an example we will take the case of a particle which, starting from rest at O, moves with a uniform speed v along OX (Fig. 24), and with a uniform acceleration a along OY, it being required to trace out the path traversed by the particle. If the particle moved with the uniform velocity v alone, the distance traversed along OX at the end of the 1st, 2nd, and 3rd second would be obtained by making t successively equal to 1, 2, 3, &c. in the equation $s=vt$, that is, they would be v, $2v$, $3v$, &c. Along OX mark the points P_1, P_2, P_3, &c., so that $\overline{OP_1}=v$, $\overline{OP_2}=2v$, &c., so that the points P_1, P_2, P_3 would represent the positions of the particle at the end of the 1st, 2nd, 3rd, &c. second if the particle were only animated with the velocity v.

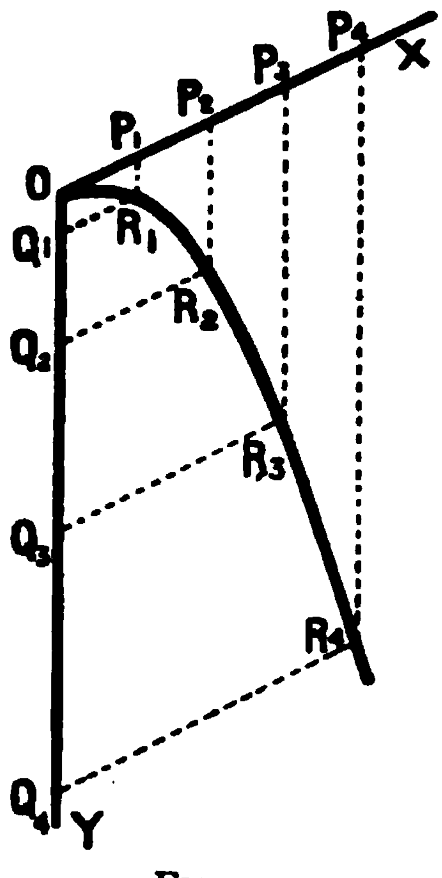


FIG. 24.

If the particle, starting from rest, were moving with the acceleration a along OY alone, the space traversed in 1, 2, 3, &c. seconds would be obtained from the equation $s=\frac{1}{2}at^2$ (§ 35) by making t successively 1, 2, 3, &c. Hence at the end of the 1st, 2nd, 3rd, &c. second the particle would be at a distance $\frac{a}{2}$, $\frac{4a}{2}$, $\frac{9a}{2}$, &c. from O, measured along OY. Let Q_1, Q_2, Q_3, &c. be points on OY, such that $\overline{OQ_1}=\frac{a}{2}$, $\overline{OQ_2}=\frac{4a}{2}$, &c., so that Q_1, Q_2, Q_3, &c. would represent the position of the particle at the end of each successive second if the acceleration a was the only motion.

Through P_1, P_2, P_3 &c. draw lines parallel to OY, and through Q_1, Q_2, Q_3, &c. draw lines parallel to OX. Then the particle at the end of the first second will have moved a distance $\overline{OP_1}$ in the direction $\overrightarrow{OX}$,

and also a distance $\overline{OQ_1}$ in the direction $\overrightarrow{OY}$, and hence will be at the point R_1. In the same way, at the end of the 2nd, 3rd, &c. second it will be at R_2, R_3, &c. The path of the particle is therefore along the curve O R_1 R_2 R_3 R_4. It will be found that this curve will be the path of an object projected *in vacuo* with a speed v in the direction $\overrightarrow{OX}$, and acted upon by gravity in the direction $\overrightarrow{OY}$.

41. Curvilinear Motion—The Hodograph.—The example considered in the last section differed from the cases previously considered in that the resultant motion, instead of taking place along a straight line, takes place along a curve. Hence not only is the speed of the particle accelerated, but the direction of motion continually changes. One of the most interesting cases of curvilinear motion is that of a particle moving in a circle, so that its speed is constant, *i.e.* it traverses equal lengths of the circumference in equal times, and the velocity changes in direction only.

The study of curvilinear motion is much simplified by the use of an auxiliary curve called the *hodograph*. The hodograph is a curve connecting the extremities of the straight lines drawn from some one given point to represent in magnitude and direction the velocity of a particle at successive instants.

Before considering the case of the hodograph of a particle moving in a curve, we will study that of a particle which moves along a polygon, each side of the polygon being traversed with uniform velocity in unit time. Let ABCDE (Fig. 25) represent the path, so that the particle takes exactly a second to travel from A to B, from B to C, from C to D, and from D to E. From any point O draw a line $\overrightarrow{OP}$ to represent the velocity with which the particle travels from A to B.

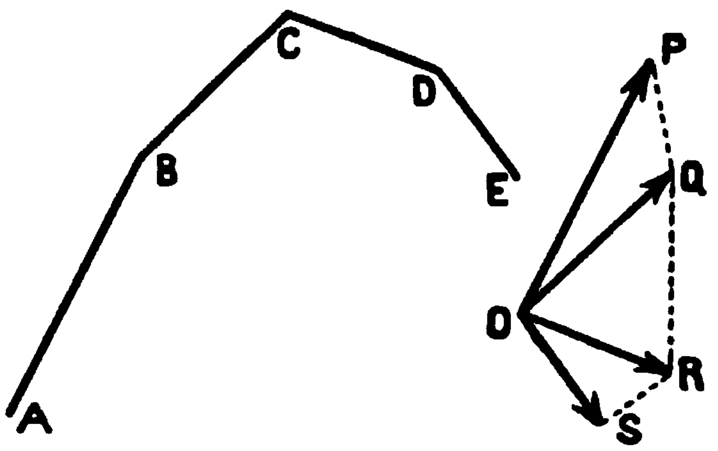


FIG. 25.

Similarly, draw $\overrightarrow{OQ}$, $\overrightarrow{OR}$, and $\overrightarrow{OS}$ to represent the velocity with which the particle travels from B to C, C to D, and D to E. Then PQRS is the hodograph of ABCDE.

In order that the particle, when it reaches B, may alter its direction of motion and travel along BC, we must compound with the velocity $\overrightarrow{OP}$, with which it reaches B, some other velocity. By the triangle of forces we see that if we compound a velocity $\overrightarrow{PQ}$ with the velocity $\overrightarrow{OP}$, the resultant velocity will be $\overrightarrow{OQ}$; that is, will be the velocity required. Hence the velocity which it is necessary to compound with the original velocity, to make the particle travel along BC with a uniform velocity so

as to reach C in one second, is $\overrightarrow{PQ}$. In the same way the velocity which has to be combined with $\overrightarrow{OQ}$ when the particle reaches C is $\overrightarrow{QR}$.

Now any curve may be considered as built up of an infinite number of very small straight lines, *i.e.* to be a polygon; hence we should expect that the results obtained above would be applicable to such a curve and its hodograph.

Suppose ABCD (Fig. 26) to be the path of a particle, and PQRS its hodograph with reference to the origin O. By this we mean that if from O lines are drawn to represent the velocity of the particle in its path ABCD in magnitude and direction at every instant of its motion the extremities of all these lines will lie on the curve PQRS.

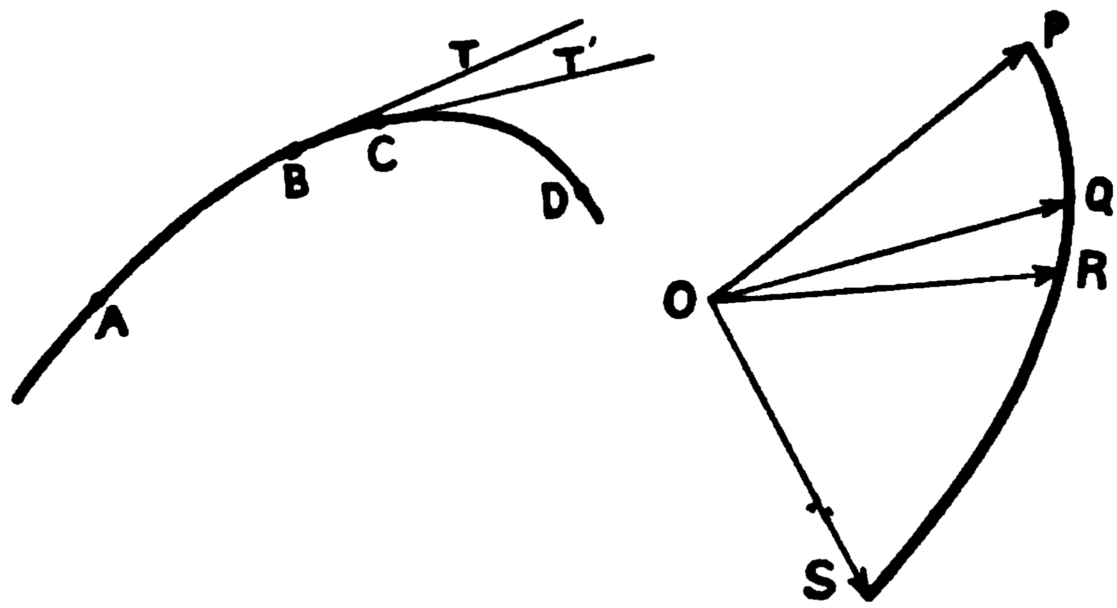


FIG. 26.

Since at any instant the direction of motion of the particle is along the tangent to its path, it follows that $\overrightarrow{OP}$ must be parallel to the tangent of the curve ABCD at A, and $\overrightarrow{OS}$ parallel to the tangent at D; also the velocity at A is equal to $\overrightarrow{OP}$, and that at D to $\overrightarrow{OS}$. To find the velocity at any other point, B, of the path, we draw a tangent BT at the given point, then in the hodograph through O draw a line OQ parallel to BT. $\overrightarrow{OQ}$ will then represent the velocity at B. In the same way $\overrightarrow{OR}$ represents the velocity at C. Suppose the particle has taken a time δt[1] to travel from B to C, then during this time the velocity has changed from $\overrightarrow{OQ}$ to $\overrightarrow{OR}$. If δt is excessively small, so that B and C, as well as Q and R, are very close together, we may regard the portion of the hodograph QR as being a straight line coincident with the tangent. Then, from what has been said with reference to the case of a polygon, the velocity which has to be compounded with $\overrightarrow{OQ}$ to give $\overrightarrow{OR}$ is represented in magnitude and direction by $\overrightarrow{QR}$.

[1] See note on p. 28.

Hence $\overrightarrow{QR}$ represents the change in velocity during the time δt. But the acceleration is defined as the change in velocity during a given time divided by the time, so that the acceleration with which the body is moving between B and C is given by $\overline{QR}/\delta t$; and, further, this acceleration acts along $\overrightarrow{QR}$, *i.e.* along the tangent at that point of the hodograph which corresponds to the point on the path at which the acceleration is being considered. If we look upon the hodograph as being traced out by a point at the same time as the particle traces out the path ABCD, then the portion QR will be traced out during the time δt that the particle spends in travelling from B to C. Hence we might call the quotient $\overline{QR}/\delta t$ the speed[1] of the tracing-point of the hodograph.

We may therefore sum up the results by saying that (1) the direction of the tangent to the hodograph at any point represents the *direction* of the acceleration of the motion of the particle at the corresponding point of its path, while (2) the speed of the tracing-point of the hodograph at any instant represents the *magnitude* of the acceleration of the particle at that instant.

42. Motion in a Circle.—The only case of curvilinear motion with which we shall deal is that of uniform motion in a circle. Since the speed is constant the hodograph is a circle, for the lines drawn from the centre O (Fig. 26) to represent the velocity of a particle moving along the circle BCD in magnitude and direction will all be of the same length, the speed being constant, and hence their extremities must all lie on a circle.

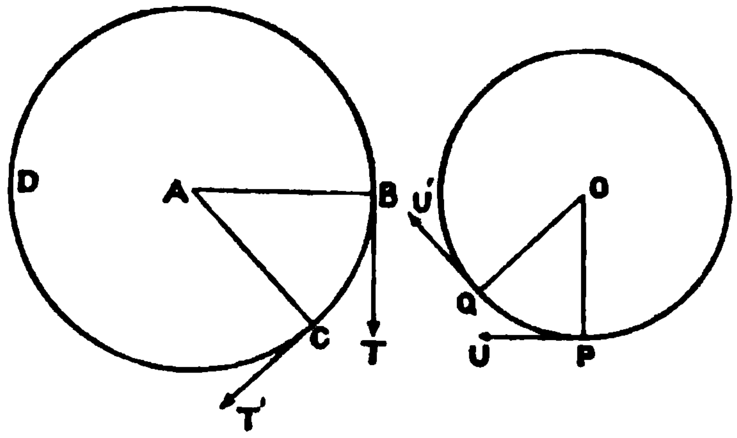


FIG. 27.

Suppose the particle to move with a uniform speed v in a circle of radius r. Then the hodograph will be a circle of radius v units of length.

Let A (Fig. 27) be the centre of the circle along which the particle moves, and B any point on the circumference. At B the particle will be moving with a speed v in the direction of the tangent BT. Hence the line $\overrightarrow{OP}$ in the hodograph represents the velocity at B, and the acceleration at B is in the direction of the tangent at P, *i.e.* in the direction PU. But since both the path and the hodograph are circles, the tangents are at right angles to the radii passing through the points of contact. Hence UP is at right angles to OP, and BT to AB. But OP and BT are by construction parallel. Hence PU is parallel to BA. The acceleration of the particle at any point is therefore along the radius of the circle at the point, the sense of the acceleration being towards the centre.

[1] The speed, if variable, is defined as in § 31.

To find the magnitude of the acceleration, we require to know the speed in the hodograph. Since the tracing-point of the hodograph will make a whole revolution in the same time that the particle describes the circle BCD, it follows that the speed in the hodograph is to the speed of the particle (v) as the radius of the hodograph is to the radius of the circle, since the circumferences of two circles are to one another as their radii, therefore the speed (u) in the hodograph is given by

$$\frac{u}{r_1}=\frac{v}{r}$$

where r_1 is the radius of the hodograph and is equal to v,

$$\therefore \qquad u=\frac{v^2}{r}.$$

Hence the magnitude of the acceleration acting on the particle is v^2/r, and the direction of the acceleration is towards the centre of the circle in which the particle is moving.

CHAPTER VI

MOTION OF A RIGID BODY

43. Definition of a Rigid Body.—A rigid body is an extended piece of matter which can move as a whole with reference to surrounding objects, but whose component particles are incapable of any displacement relative one to the other. Hence a rigid body is incapable of having a strain (§ 122) impressed upon it. Although a rigid body is an ideal which cannot be realised in practice, the consideration of the dynamics of a rigid body is useful as an introduction to the study of the more complex problems which arise when we have to deal with such substances as exist in nature, besides which, in many problems, such bodies as steel and glass may be taken as rigid.

44. Motion of a Rigid Body.—Any displacement of a rigid body can be produced by a pure translation of the body, and a pure rotation of the body round a certain point called the centre of figure[1] of the body. Thus to consider the case of the displacement of a straight line AB (Fig. 28) in a plane, from the position AB to the position A′B′. In the case of a straight line the centre of figure is at the middle point. Hence in the first position C is the centre of figure, and in the second C′. The line AB may be displaced to the position A″B″ by a pure translation (§ 29), since all the particles will move along equal parallel straight lines. Then it can be rotated about C′ into the required position A′B′. In this case all the particles will move in circles with C′ as a centre. The motion of a rigid body may thus be resolved into a motion of translation obeying the laws considered in Chapter V., and a motion of rotation.

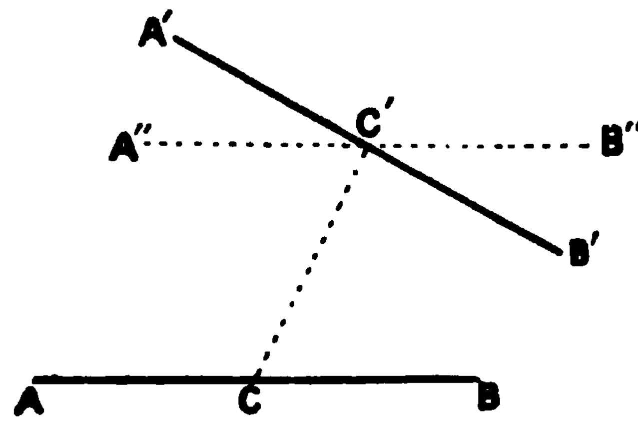


FIG. 28.

45. Motion of Rotation.—When a rigid body moves so that the paths of all the particles of which it may be regarded as built up are circles having their centres on a fixed line, called the axis of rotation, the body is said to undergo a pure rotation about this axis.

Since the distances of the different particles from the axis of rotation

[1] When we come to consider the action of gravity on bodies, we shall find that the centre of figure is what is better known as the centre of gravity.

may not all be the same, but as they are all rigidly attached together they must complete a revolution in exactly the same time, it follows that the speed of the different particles is different. Thus in the case of a fly-wheel, consider the motion of two points, one on a spoke at a distance of 50 centimetres from the axis of rotation (the axle), and the other on the rim, say at a distance of 100 centimetres from the axis. Suppose the fly-wheel to make one turn per second, then the first particle will in one second traverse the circumference of a circle of 50 centimetres radius, *i.e.* will travel through 100π centimetres, and hence its speed will be 100π centimetres per second. The speed of the other particle will in the same way be 200π centimetres per second.

The velocity of rotation of a body cannot therefore be measured by the speed of any unspecified particle, but is measured either by the number of turns made in a given time divided by that time, *i.e.* the number of turns per second, or by the speed of a particle at unit distance from the axis of rotation, this speed being called the *angular velocity* of the body.

Suppose the angular velocity of a body is ω, that is the linear speed of a particle at unit distance from the axis is ω, so that the space passed over by such a particle in a second is ω units of length. Now the length of the arc of the circle, along which the particle travels, passed over in one second being ω, the angle subtended by this arc at the centre of the circle is $\omega/1$ in circular measure (§ 14), since the radius of the circle is unity. Since the *angle* swept out by the radius joining any particle of the body to the axis of motion is the same for all particles in the body, we may say that the angular velocity ω of a body represents the angle (measured in *circular measure*) through which the body turns in one second.

If the rotating body turns through an angle ω in one second, then the space traversed by a particle at a distance r from the axis of motion is ωr, since this is the length of an arc of a circle of radius r that subtends an angle ω at the centre.

The relation between the angular velocity of a body and the number of turns per second can be obtained as follows: Let the body make n turns per second, then since in one complete turn the angle turned through is 2π (§ 14), the angle turned through in one second, or the angular velocity, is $2\pi n$. Since the body makes n turns per second, the time it takes to make one turn is $\frac{1}{n} = T$ say, and the angular velocity $= 2\pi/T$.

A rotation is said to be positive if, when looking along the axis in the positive direction, it appears to take place in the opposite direction to that of the hands of a watch; while if it takes place in the same direction it is negative.

As in the case of linear velocity, angular velocity may be uniform or variable. In the case of variable angular velocity, the change of angular

velocity in a given time divided by that time is called the angular acceleration. Angular acceleration may be uniform or variable. Since the angular velocity is measured by the angle swept out in unit time, and the angle is measured by the arc swept out divided by the radius (*i.e.* angular velocity = (arc ÷ radius)/time), the dimensions of angular velocity are given by $[\Omega]=[(L\div L)\div T]=[T^{-1}]$. The truth of this result will be apparent if we remember that the angular velocity does not depend upon the distance of the point considered from the axis of rotation, *i.e.* the size of the circle described, but simply on the time taken to go round this circle. The dimensions of angular accelerations are $[T^{-2}]$.

46. Composition and Resolution of Rotations. — In order to completely define a linear velocity we have seen that we require to know the direction of the velocity, and also the magnitude of the velocity. In the same way, to define a rotation we require to know the axis about which the rotation takes place and the magnitude of the rotation, *i.e.* the angular velocity. Rotations may be compounded or resolved in a manner analogous to that employed in the case of motions of translation. Thus, if a rigid body has simultaneously applied to it two rotations about different axes passing through a fixed point, the resultant motion consists in a rotation about a single axis passing through this point. The direction of the axis of this resultant rotation and its magnitude can be obtained by drawing through any point (O) two straight lines parallel to the axes of the two given rotations, and equal in length to the angular velocities, and completing the parallelogram. The diagonal drawn through O will then represent the axis of the resultant rotation in direction and the angular velocity in magnitude. In the same way we may resolve any rotation into two or more component rotations round any given axes by employing the methods used in § 38.

47. Degrees of Freedom of a Body. — In § 26 we saw that in order to define the position of a point in space we require to know its distance from *three* fixed planes, no two of which are parallel. Again, every possible way in which a particle can move may be resolved along *three* mutually perpendicular lines or axes, and the resultant of these three component displacements will be equal to the original displacement. Hence, when the motion of a *particle* is unrestricted by any conditions it is said to have *three degrees of freedom*.

Next, suppose the particle is constrained to remain always in contact with a given surface. It will then only possess two degrees of freedom, for we may take the normal to the surface as one of the three mutually perpendicular directions along which the motion of the particle may be resolved. Since no displacement of the particle can take place along this direction, or the particle would not remain in contact with the surface, the only remaining possible independent displacements are at right angles to one another in the tangent plane to the surface.

If the particle is constrained to move so as to remain in contact

simultaneously with each of two surfaces, it has only one degree of freedom. For it is constrained to move along the line which is common to both surfaces, *i.e.* the line of intersection of the surfaces, and if a point moves on a line, it can, at each point, only move in one direction, which is that of the tangent to the line at the given point.

In the case of a perfectly unconstrained extended rigid body we have *six* degrees of freedom, for in addition to the three independent translations possessed by a material particle, an extended body is also capable of three independent *rotations* about three mutually perpendicular axes.

If such a body have one point fixed, then it loses all its possible motions of translation (*i.e.* three degrees of freedom), and is only capable of three rotations about axes passing through the fixed point.

If two points in the body are fixed, then in addition it loses two rotations, since the only motion of which it is now capable is a rotation about an axis passing through the two fixed points.

If three points, not all in a straight line, are fixed the body loses all its degrees of freedom and is fixed. We thus see that if we are given the positions of any three points (which are not all in the same straight line) of a body the position of the body is completely defined.

48. Geometrical Clamps and Slides.—An interesting practical application to the construction of instruments of the principles enunciated in the previous section has been made by Lord Kelvin, the importance of which is only slowly making itself felt even in the construction of scientific instruments.

Suppose that an instrument, standing on three legs, rests on a horizontal plane, then it has lost one translation, namely, that perpendicular to the plane, and also two rotations, the only rotation left being about an axis perpendicular to the plane. If, further, one of the legs rests in a conical hollow A (Fig. 29) the instrument has lost all its translations in addition to two rotations. If, therefore, we prevent the only degree of freedom left, *i.e.* the rotation about an axis perpendicular to the plane passing through the conical hole, the instrument will be fixed in position. We may prevent this rotation by allowing one of the two legs which now rest on the smooth surface of the table to rest in a V-shaped groove B pointing towards the conical hole, the third leg resting on the plane surface at C. This arrangement, called the hole, slot, and plane, forms a geometrical clamp and enables us to remove the instrument and yet replace it in exactly the same position as it before occupied, and this without any complicated arrangement of screws and clamps.

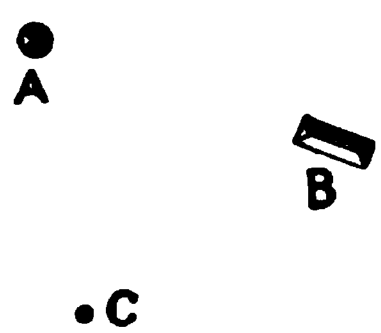
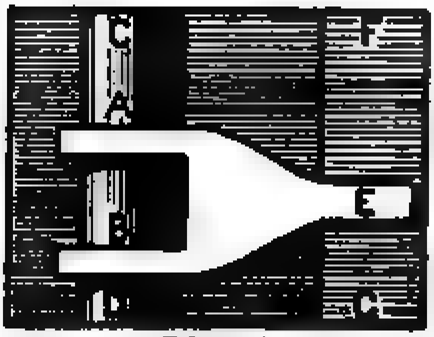


FIG. 29.

As an example of a geometrical slide we may take the case illustrated

in Fig. 30. Here two V's (A and B) attached to the body rest on the surface of a cylinder (C D). Each V touches the cylinder in two points, and therefore the only motions left to the body are (1) a translation parallel to the axis of the cylinder, and (2) a rotation about the axis of the cylinder. Hence, if this rotation is restrained by a point E being kept pressing on a plane FG parallel to the axis of the cylinder, the only motion possible is a translation parallel to this axis. The peculiarity of a geometrical slide is that it enables us to obtain the desired constrained motion without any possible rattle or looseness, which in the form of mechanical slide ordinarily used is certain to be found sooner or later, owing to the wear of the parts, even if it does not exist originally when the slide is new. Want of space will not permit of our following this subject any further, but those who are interested in mechanical design will find that a careful study of the properties of geometrical clamps and slides fully repays the time spent, for no amount of good workmanship can compensate for bad design.



PLAN

FIG. 30.

CHAPTER VII

PERIODIC MOTION

49. Definition of Periodic Motion.—We have hitherto considered the motion of a body to be either uniform or to vary continuously. There is, however, a most important kind of motion in which the body goes through the same series of movements at regularly recurring intervals. This kind of movement is called periodic. If in addition to being periodic the motion is continually being reversed in direction, it is said to be vibratory or oscillatory. The motion of the earth with reference to the sun is periodic, since although the velocity is not uniform, yet at regularly recurring intervals (the year) the velocity regains the same value. The motion of a pendulum or that of the prong of a tuning-fork are examples of oscillatory or vibratory motion.

50. Simple Harmonic Motion.—If we look at a particle which is moving uniformly in a circle (MQN, Fig. 31) from some point P in the plane in which the particle is moving, and from some distance off, the particle will appear to move backwards and forwards along a straight line MN. While the particle moves over the arc NQ′M, it appears to be moving from left to right; it will then appear momentarily to come to rest as it reaches the

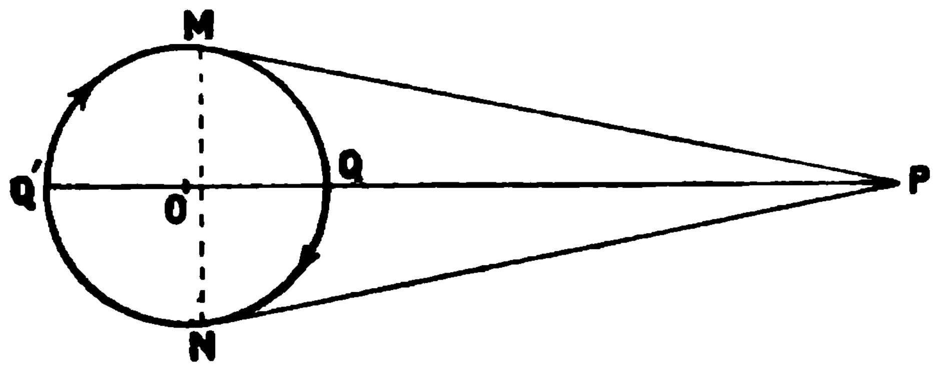


FIG. 31.

extreme limit of its path to the right; then it will start moving to the left, at first slowly, but with increasing speed till it appears in line with the centre O of the circle on which it is actually moving. The speed will now appear to diminish till the particle comes to rest for a moment at the extreme left-hand limit N of its path. It will then appear to travel to the right, the speed increasing till it passes the centre, then diminishing till it again reaches its extreme right-hand position. Now all the time the particle is in reality moving at a uniform speed in a circle, and when it appears at rest it is only moving either directly towards or away from our eye. Thus when the particle is at M it is moving momentarily along the line MP, which is a tangent to the

circle passing through P; and hence, since we should be unable from P to detect this movement in the line of sight, it would appear at rest. In the same way at N it is really moving in the direction PN, although it appears stationary. When it is at Q or Q′ it is moving at right angles to the line of sight, and hence appears to move with the greatest speed; at Q from right to left and at Q′ from left to right. If the point P is at a very great distance from the circle, the tangents PM and PN will be parallel, and NM will be a diameter of the circle.

It will save circumlocution if we suppose that a second particle moves to and fro along the diameter MN (Fig. 32) in such a way that it always appears, when viewed from a great distance, in a line with the first particle, which is moving in the circle NQM. This means that this second particle is at any moment at the foot of the perpendicular drawn from the position of the first particle to the diameter MN. Thus when the particle which moves in the circle is at the points A, B, C, &c., the other particle will be at *a*, *b*, *c*, &c. Suppose the one particle makes a complete revolution in two seconds, so that it traverses the semi-circumference NQM in one second. Since by supposition it moves with uniform angular velocity (§ 45), if we divide the semicircle NQM into ten equal arcs, NA, AB, BC, &c., each of these arcs will be traversed in a tenth of a second. If then we draw perpendiculars to the diameter MN through all these points, the feet of these perpendiculars, *i.e.* the points *a*, *b*, *c*, *d*, &c., will represent the positions of the particle which moves along the diameter at $\frac{1}{10}$, $\frac{2}{10}$, $\frac{3}{10}$, &c. second after it leaves N. In order to find the distances O*a*, O*b*, O*c*, &c., of the particle from the centre O, we join the points A, B, C, &c., to the centre. Then, if *a* is the radius of the circle, we have the following relations—

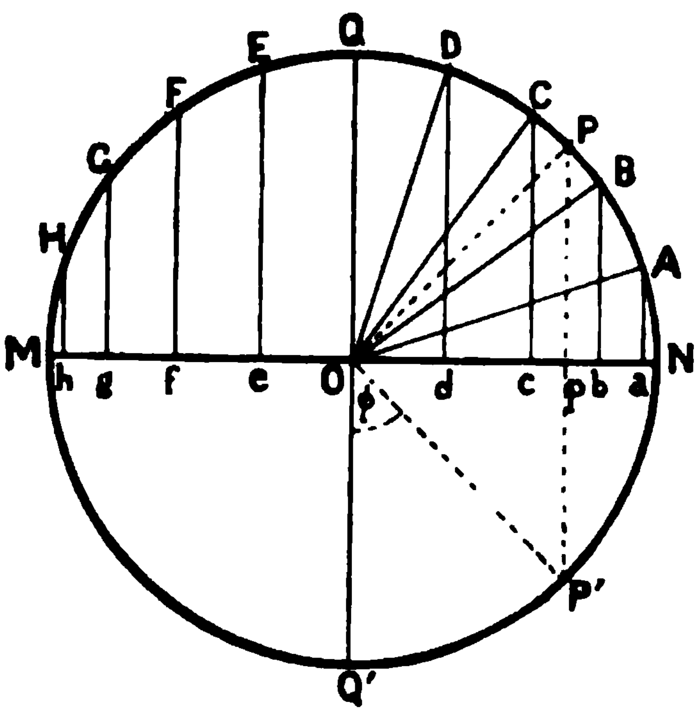


FIG. 32.

$$Oa = a \cos AON,$$
$$Ob = a \cos BON = a \cos (2.AON),$$
$$Oc = a \cos CON = a \cos (3.AON), \&c.$$

If ω is the angular velocity of the particle moving in the circle, the angle swept out by the radius joining the particle to the centre in one second is ω, and hence, since during a complete revolution the angle swept out is 2π, the particle will make a complete revolution in $\frac{2\pi}{\omega}$ seconds. This will therefore represent the time which elapses between

two consecutive passages ***in the same direction*** of the particle that moves along MN through any given point in its path, and is called the ***periodic time***, or ***period*** of the oscillatory motion of the particle.

If we start measuring our time from the instant when both the particles are at N, and call the distance of the vibrating particle from the centre, O, x, x being positive when the particle is to the right of O and negative when it is to the left of O, then at a time O, $x=a$; at a time $\frac{2\pi}{\omega}/2$, $x=-a$, since the particle will now be at M. If P is the position of the particle moving in the circle, and p that of the vibrating particle at the time t, we have

$$x=a\cos \text{PON}.$$

Now PON is the angle swept out by the particle moving in the circle in the time t. Hence the angle PON$=\omega t$, for (§ 45) ω is the angle swept out in unit time, and therefore ωt is the angle swept out in the time t. Therefore

$$x=a\cos\omega t \quad . \quad . \quad . \quad . \quad (1).$$

When $t=0$, $\cos\omega t=1$, and hence $x=a$, a result already obtained. When $t=\pi/\omega$, *i.e.* half a period later, $x=a\cos\omega\times\frac{\pi}{\omega}=a\cos\pi=-a$, that is, the particle is at M. When $t=2\pi/\omega$, *i.e.* a whole period after the start, $x=a\cos\omega.\frac{2\pi}{\omega}=a\cos 2\pi=a$, and the particle is back at N. When

$$t=\pi/2\omega, \text{ or } 3\pi/2\omega$$

$$x=a\cos\omega.\frac{\pi}{2\omega}=a\cos\frac{\pi}{2}=0$$

or

$$=a\cos\omega.\frac{3\pi}{2\omega}=a\cos\frac{3\pi}{2}=0,$$

and hence the particle is at O. We thus see how, for some easily recognised positions, equation (1) gives the position of the vibrating particle in terms of the time.

A particle that moves to and fro along a line, such as MN, in the way the particle above considered does, and so that its position with reference to O, the middle point of MN, is always given by equation (1) above, is said to move with a ***simple harmonic motion***. Such a movement we shall in future indicate by the initials S.H.M. The reason for giving it this name will appear when we come to study acoustics, for it will be found that most of the movements with which we are there concerned are S.H.M.'s.

The maximum distance from the median position O, attained by a particle which is executing a S.H.M., is called the ***amplitude*** (a). It will be noticed that the amplitude (a) is equal to the radius of the circle MQN (Fig. 32) used to define the S.H.M. This circle is called the ***circle of reference***.

The angle which the radius, passing through the particle in the circle

of reference, makes with the positive direction of the path of the vibrating particle, *i.e.* with ON (Fig. 32), is called the *phase* of the particle executing S.H.M. at the corresponding instant. Thus when the vibrating particle is at N the phase is zero, when at P the phase is equal to the angle PON. We have seen that the angle PON is equal to ωt where ω is the angular velocity in the circle of reference, and t is the time counted from the passage of the vibrating particle through its position of maximum positive elongation.

Instead of starting to measure the time, or what comes to the same thing, the phase of the vibration, from the instant when the particle passes through the position of maximum positive elongation, it is often more convenient to start at the instant when the particle passes through its mean position, O, when moving in the positive direction, so that both the time and the displacement are measured from the conditions of the particle when passing through O.

Suppose we measure the phase from the line OQ′ (Fig. 32) when the particle is passing through O in the direction from M to N. Let P′ be the position of the particle in the circle of reference at a time t after the start from Q′. Then Op is the displacement of the particle which is executing a S.H.M. along MN at the time t. Now

$$\begin{aligned} \mathrm{O}p &= a \cos \mathrm{P'O}p \\ &= a \sin \mathrm{P'OQ'} \\ &= a \sin \phi, \end{aligned}$$

where ϕ is the phase measured from OQ′. As before, since we are now measuring time from the instant the particle in the circle of reference passes through Q′,

$$\phi = \omega t.$$

Hence the displacement x is given by the equation

$$x = a \sin \omega t \quad . \quad . \qquad (2).$$

Thus the form of the expression for determining the displacement depends on whether we start measuring our time from the instant when the particle is at its extreme elongation or when it is passing through its mean position. The motion represented by the two expressions is, however, the same, if the quantities a and ω have the same values in the two cases.

If T is the periodic time of the S H.M., then, since the particle moving in the reference circle with uniform angular velocity ω will complete a whole rotation in a time $\frac{2\pi}{\omega}$, we have

$$T = \frac{2\pi}{\omega}$$

$$\text{or } \omega = \frac{2\pi}{T}.$$

The number of periods per second, or the number of complete, *i.e.* to-and-fro oscillations per second ; or, what amounts to the same thing, the

number of revolutions in the circle of reference per second is called the *frequency* of the S.H.M. Hence if N is the frequency,

$$N = \frac{1}{T} = \frac{\omega}{2\pi}.$$

51. Velocity and Acceleration in S.H.M.—The velocity of the particle in the circle of reference at a point P (Fig. 33) is ωa (§ 45) along the tangent PT. We may resolve this velocity along directions PB and PA, parallel and perpendicular to MN.

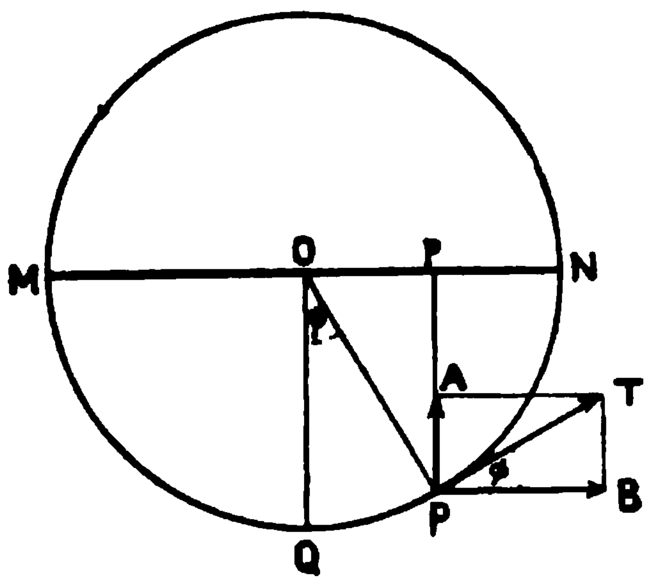


FIG. 33.

Let us measure the phase ϕ from the instant when the particle passes through the mean position from left to right. Now PT (the tangent) is perpendicular to OP (the radius), and PB is perpendicular to OQ; therefore the angle TPB is equal to the angle QOP or ϕ. Hence the component along PB of the velocity ωa is

$$\omega a \cos \phi.$$

The component of the velocity at right angles to MN cannot affect the space passed over parallel to MN, so that if we only consider the resolved part of the motion of P parallel to MN, the velocity at any point would be $\omega a \cos \phi$. But this resolved part of P's motion is the same as the motion of the particle which executes a S.H.M. along MN, and of which P moves in the circle of reference. Hence the velocity $\dot{x}$[1] of the particle, moving with S.H.M., when at P is given by the equation

$$\dot{x} = \omega a \cos \phi.$$

But $\cos \phi = \overline{Pp}/\overline{OP} = \sqrt{(\overline{OP}^2 - \overline{Op}^2)}/\overline{OP} = \sqrt{(a^2 - x^2)}/a$, where x is the displacement of the particle from its mean position. Hence

$$\dot{x} = \omega \sqrt{(a^2 - x^2)}.$$

When $x = +a$ or $-a$, that is, when the particle is at the points N or M, the velocity $\dot{x}$ is zero. This is also evident, since at these points the particle in the circle of reference is moving at right angles to MN. The maximum value of $\dot{x}$ occurs when $x = 0$, and is $a\omega$, as is evident, since at O the velocity of the particle executing the S.H.M. is the same as that of the particle in the circle of reference.

The acceleration of a point moving in the reference circle with a linear speed v is constant, directed towards the centre and equal to v^2/a (§ 42). But $v = a\omega$, hence the acceleration is equal to $a\omega^2$.

This acceleration may be resolved parallel and perpendicular to

[1] The symbol $\dot{x}$ is used to denote the velocity of the particle when its displacement is x.

MN, and the component at any point parallel to MN will be the acceleration in the S.H.M. at the corresponding point. The component along Pq (Fig. 34) will be $a\omega^2$ cos OP$q = a\omega^2$ sin ϕ, hence this is the acceleration in the S.H.M. at p.

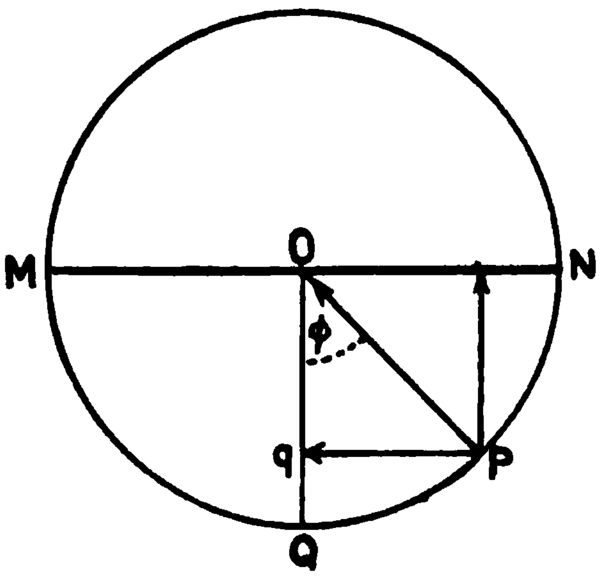


FIG. 34.

We found in § 50 that the displacement (*i.e.* Op) at a given point p was equal to a sin ϕ. Substituting this value for a sin ϕ in the expression for the acceleration, we find that when the distance of the particle from the median position, or the displacement, is equal to x the acceleration $\ddot{x}$[1] is given by

$$\ddot{x} = \omega^2 x.$$

This shows, since ω is constant, that the acceleration is simply proportional to the displacement, and since it always acts towards the centre O, if the particle executing S.H.M. is moving away from O the acceleration is negative, *i.e.* is a retardation; and if the particle is moving towards O the acceleration is positive, and hence the velocity is increasing. This agrees with what we previously found, namely, that the velocity is a maximum as the particle passes through its median position.

52. Harmonic Curve.—Suppose we have a particle executing a S.H.M. along MN (Fig. 35) of period T and NQM is the circle of reference. Then, supposing the particle to start from its median position,

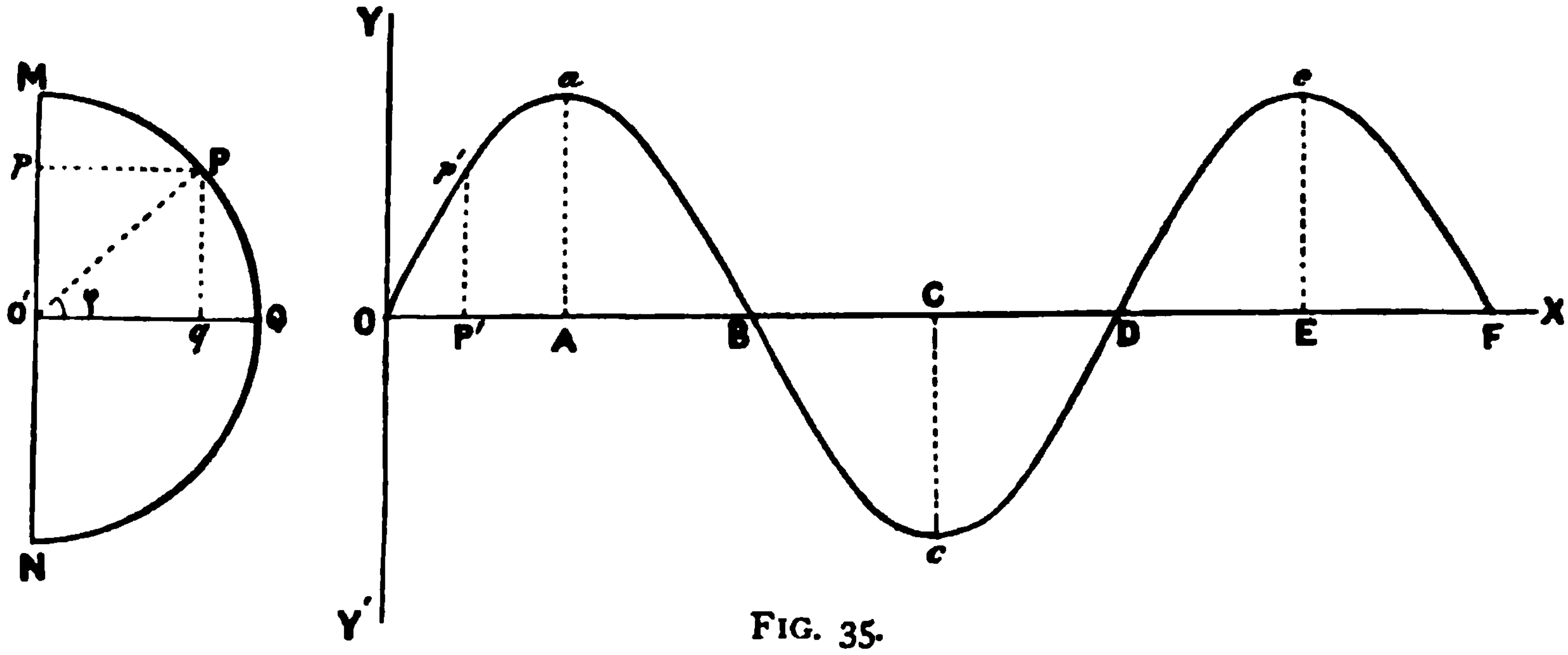


FIG. 35.

let us draw a diagram in which the time is measured along OX, and the displacement of the particle from its *median* position along OY. Since the particle is supposed to start from o', the displacement at the time O

[1] The symbol $\ddot{x}$ is used to denote the acceleration with which the particle is moving when the displacement is x.

will be zero. At a time $T/2$ later the particle will have again come back to o', and the displacement will again be zero. Hence, if OB represents the time $T/2$, the displacement at this point will be zero. A point (A) half-way between O and B will represent a time $T/4$, and the displacement at this time is o'M, since, starting from Q in the circle of reference, the particle in $\frac{1}{4}$ of the time (T) taken to make a whole revolution will have traversed $\frac{1}{4}$ of the circumference, and will therefore be at M. Hence at A we erect a perpendicular Aa equal to o'M. At the point C, which corresponds to a time $3T/4$, the particle will be at N, and hence the displacement is o'N, equal in magnitude to the displacement at M, but opposite in direction. We therefore draw Cc in the opposite direction to Aa, and make it equal to o'N or o'M. At a time T the particle will have completed a vibration and will be back at o', so that the displacement at D, where OD$=T$, is zero. If we drew ordinates to represent the displacement at intermediate times, and then drew a line through the extremities to these ordinates, we should obtain the wavy curve OaBcDeF. This curve, which represents the relation between the displacement of the particle which is executing S.H.M. and the time, is called the ***harmonic curve***.

This curve is of great interest from its bearing on many physical problems, so that it will repay us if we investigate a few of its properties. Suppose that at the time t the particle executing S.H.M. is at p, so that the corresponding points in the circle of reference and on the harmonic curve are P and p' respectively. Then the displacement is $o'p$ or P$'p'$. Now $o'p=$ P$q=o'$P sin Qo'P$=a \sin \phi$, where a is the amplitude and ϕ is the angle Qo'P[1]. Hence P$'p'=a \sin \phi$. Now $\phi=\omega t$, therefore P$'p'=a \sin \omega t$.

In the harmonic curve the abscissæ represent the time, so that OP$'=t$. Hence if we call the abscissa of any point on the harmonic curve x and the *corresponding* ordinate y, since $y=$ P$'p'$ and $x=$ OP$'=t$, we get

$$y=a \sin \omega x.$$

The harmonic curve is therefore sometimes called the *curve of sines*, being such that the ordinate at any point is proportional to the sine of an angle which is itself proportional to the abscissa.

The actual form of the curve depends on the amplitude a and on the angular velocity ω in the reference circle of the S.H.M., or, since $\omega=2\pi/T$, on the periodic time of the S.H.M.

53. Composition of Simple Harmonic Motions.—S.H.M.'s, like any other form of motion, can be compounded, and the composition can in general be best effected by a geometrical method by means of their circles of reference.

The simplest case is that of two S.H.M.'s in the same direction, of

[1] ϕ is the *phase* of the vibration measured from the instant when the particle is moving through its mean position (see § 50).

the same period, and equal amplitude. If the phases are the same, the resultant will be a S.H.M. of the same period and phase as the constituent motions, but of double the amplitude. The S.H.M.'s being in the same phase, the resultant displacement at any moment will be the sum of the two component displacements, *i.e.* since these are equal, twice either of them. This double displacement is exactly what would occur if the radius of the circle of reference were twice as great as in the case of the component motions, for the triangles $PO p$ and $QO q$ (Fig. 36) are similar, and hence

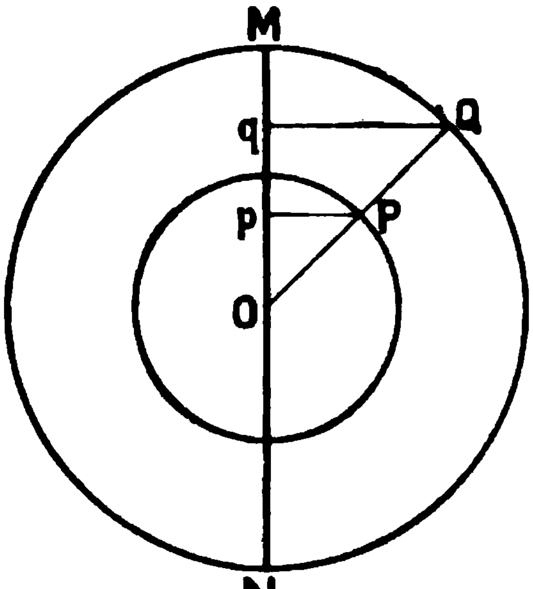


FIG. 36.

$$\frac{\overline{qO}}{\overline{pO}}=\frac{\overline{QO}}{\overline{PO}}=2.$$

If the phases of the S.H.M.'s differ by half a period (180°) they will exactly neutralise each other and will produce rest, the displacement at any instant due to one motion being exactly equal and opposite to that due to the other.

Next, suppose the two S.H.M.'s are at right angles, but of the same period, of equal amplitudes, and in the same phase. Let one motion take place along KL (Fig. 37) and the other along MN, and let KMLN be the common circle of reference. Divide the circumference of this circle into an even number of equal parts, say twelve, and through these points draw lines parallel to KL and MN as in Fig. 37. As the S.H.M.'s are in the same phase, the extreme positive elongation will occur at the same instant in each. Hence if, as is usual, we consider from O to L to be the positive direction for one motion, and from OM that for the other, L will be the position of maximum positive elongation for one and M for the other.

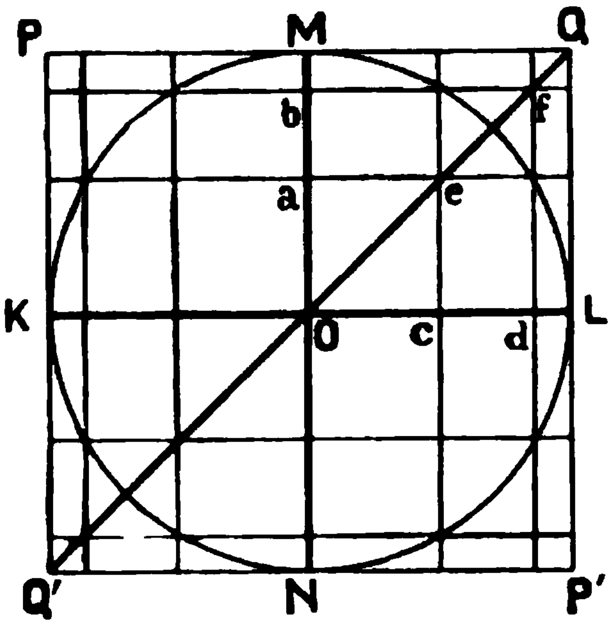


FIG. 37.

Starting then from the instant when both the motions are passing through the position of rest O, and the positive displacement is increasing, the points a, b, M will represent the displacements at times $T/12$, $2T/12$, and $3T/12$ due to the S.H.M. along NM, while the points c, d, L will represent the displacements at the same instants due to the motion along KL. Hence the actual displacement of a particle which is moving with the two S.H.M.'s will be O, e, f, Q, &c. The resultant motion is therefore along the straight line QQ′, which is inclined at 45° to the

directions of the two S.H.M.'s. For the amplitude and frequency of the two motions being the same, the displacements Oc, Od, and OL are equal to the displacements Oa, Ob, and OM respectively.

Since

$$OQ = OL/\cos 45^\circ\ ;\ Of = Od/\cos 45^\circ\ ;\ Oe = Oc/\cos 45^\circ,$$

it follows that the resultant displacement is always equal to the corresponding displacement in one of the component S.H.M.'s divided by the cosine of 45° or $1/\sqrt{2}$. Now the displacement x along KL can be represented by the equation $x = a \sin \omega t$ (§ 50). Hence the resultant displacement along QQ′ can be represented by

$$R = \sqrt{2}.a \sin \omega t.$$

This represents a S.H.M. of which the periodic time is the same as that of the two components (ω being the same for all three), and of which the amplitude is $\sqrt{2}.a$.

If the two S.H.M.'s, instead of being in the same phase, differ in phase by half a period, or 180°, then the resultant motion will be as S.H.M. along PP′ of amplitude $\sqrt{2}.a$.

If the two components differ in phase by 90°, or a quarter period, when one S.H.M. is at its extreme elongation the other will be passing through its position of rest. Suppose that when the moving particle is at the point of extreme positive elongation (M), Fig. 38, as far as its motion along MN is concerned it is passing through O from left to right, owing to the motion along KL. Then at successive intervals of $T/12$ it will be displaced to b, a, O respectively with reference to one motion, and to c, d, L with reference to the other; and hence its resultant position will be e, f, L, &c. The resultant motion will thus be uniform motion in the circle of reference in the direction MLNK. If, however, when the particle is displaced to L by the horizontal motion, it is passing through O in the direction NOM, the resultant motion will be in the circle of reference but in the direction LM KN.

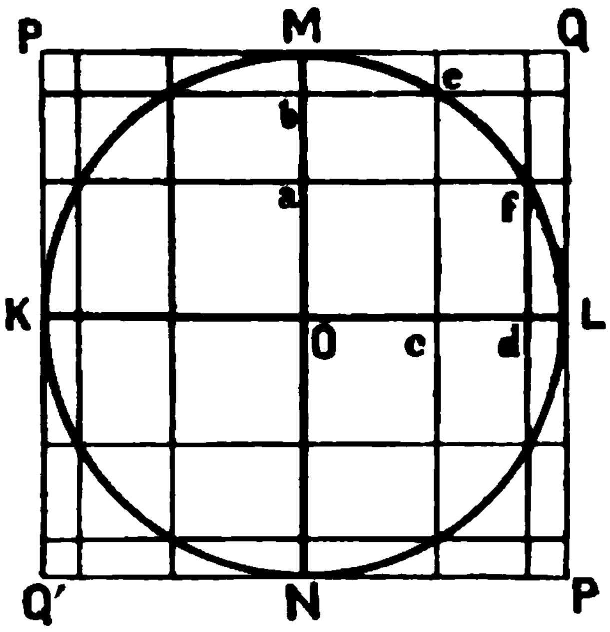


FIG. 38.

For any other difference of phase the resultant motion will be in an ellipse, which will touch all four sides of the square PQP′Q′.

When either the amplitudes or periods of the two component S.H.M.'s are different, we cannot use the same circle of reference for the two motions. Suppose the period of the vertical S.H.M. is 2/3 that of the horizontal S.H.M., and the amplitude of the vertical motion is 2/3 that of

the horizontal, the phases being the same. Let M′AN′ (Fig. 39) and K′BL′ be the two circles of reference, the diameter M′N′ being 2/3 the diameter K′L′, since the amplitude of the motion along MN is 2/3 of that along KL. The circumferences of these two circles must next be divided into parts that are traversed by the tracing-points in equal times. It is convenient to divide the quadrant into a whole number of parts in each case, hence in the above example we divide the circle M′AN′ into eight parts, and the circle K′BL′ into twelve parts. The period of the motion along MN being 2/3 of that along KL, the tracing-point of the circle M′AN′ will traverse the circumference, while the tracing-point of the circle K′BL′ traverses 2/3 of the circumference. Hence the tracing-point in M′AN′ will traverse 1/8 of the circumference in the same time that the tracing-point in K′BL′ traverses 1/8 of 2/3 or 1/12 of the circumference; and hence we have divided the circles so that the arcs will be traversed in equal times. The phase of the motions being the same, the two extreme positive elongations occur simultaneously, and the particle starts at P. At the end of the interval chosen for subdividing the circles it has moved down to *a*, and horizontally to *d*, and hence its position is at Q. At the end of the next interval it has moved downwards to *b*, and horizontally to *e*; it is therefore at R. Similarly it travels to S and T. At T the particle has reached its extreme elongation in the vertical direction, and hence it now begins to move upwards, and during the next interval it reaches *c*. It continues, however, to move to the left in the horizontal direction, and at the end of the interval is displaced to *h*. The actual position is thus U. In a similar manner the position at the ends of the remaining intervals can be found, and the path will be given by the line PRTKRP′. When the particle reaches the point P′, which it does after one complete period of the slower vibration (*i.e.* the horizontal) and one and a half periods of the faster, it will retrace the path, returning to P after two complete periods of the slower S.H.M. and three of the faster.

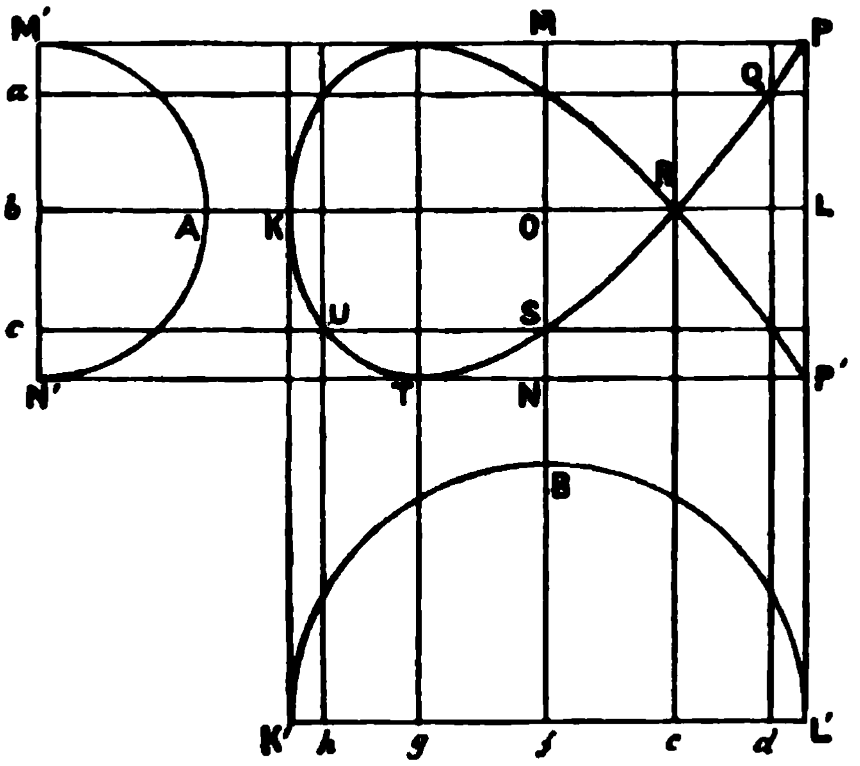


FIG. 39.

If the phases of the two components are not the same, the resultant motion would be different; the method of drawing the curves is, however, the same as in the above example. Some of the figures obtained are given in Fig. 40, where the phase of the vertical S.H.M. is increased by

45° between each figure and the next. In Fig. 41 another series of curves is given, in which the periods of the component S.H.M.'s are as 1 : 2, the amplitudes being the same. In this case the phase of the S.H.M. of shorter period is advanced by 30° between each figure and the next.

The above are all examples of the composition of two S.H.M.'s, the periods of which are commensurate; that is, the ratio of the periods is expressed by *simple* whole numbers, so that, after a comparatively short time, equal to the least common measure of the periodic times, the

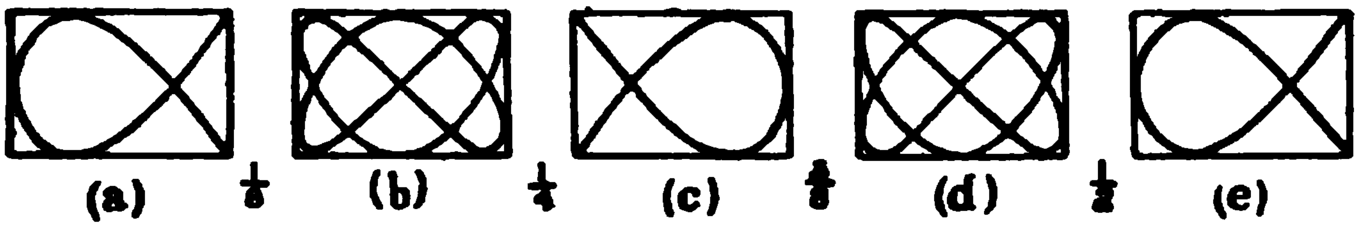


FIG. 40.

particle will come back to its starting-point and the curve will then be retraced. If, however, the periods are not commensurable, the particle will not come back to its starting-point till after an infinite number of complete periods; that is, not at all.

There is one case which is of considerable interest, that is, when the periods can very nearly be represented by two simple whole numbers. If, for instance, the periods are as 2 : 1, then, as we have seen, we get a series of curves according to the difference in phase between the component motions; in each case, however, the curve is constant in form. Suppose

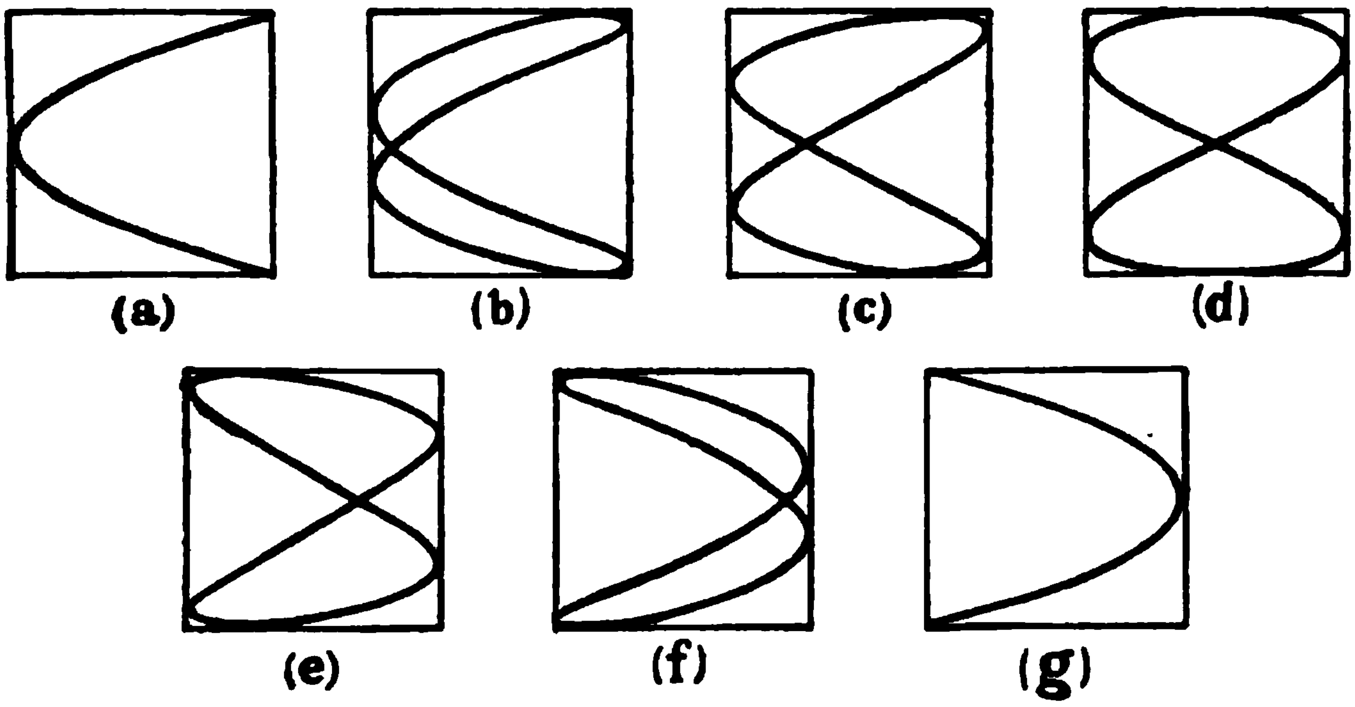


FIG. 41.

now the periods, instead of being as 2 : 1, are as 201 : 100, and that the S.H.M.'s start in the same phase, then the path of the particle will be very nearly like (*a*), Fig. 41. However, when the slower S.H.M. has completed one vibration, the other, instead of having exactly completed two vibrations, will have completed two whole vibrations, together with

1/100 of another; it will thus have gained in phase on the other by 1/100 of a period or $360/100 = 3^\circ.6$. This gain will continue till after eight periods of the slower vibrations, the difference in phase will amount to $28^\circ.8$, and hence the curve traced out will resemble (*b*), Fig. 41. The difference in phase will continue to increase, and so by a continuous modification the curve will pass in succession through all the forms shown in Fig. 41, first from (*a*) to (*g*), and then back from (*g*) to (*a*). For after 100 periods of the slower vibration, the quicker will have made a whole vibration more than it would have made if the ratio of the periods had been exactly 2 : 1, and for an instant the curve will again take the form of (*a*), Fig. 41, and will then go through the whole series again.

54. Composition of Two Simple Harmonic Motions in the Same Direction.—In the previous section we have dealt with the composition of two S.H.M.'s, when the directions of motion are at right angles. We have now to consider the composition of two S.H.M.'s when the directions of motion are along the same straight line, the simplest case of which, namely when the amplitude and phases of the S.H.M.'s were the same, we considered on p. 58. The simplest method for effecting this composition is by means of the harmonic curve (§ 52).

Let ABCDE (Fig. 42) be the harmonic curve corresponding to one

FIG. 42.

S.H.M., so that AM represents the amplitude and MN the period, and let *abcde* represent another S.H.M. of amplitude *a*M and period M*f*, which starts in the same phase as the other. Then the resultant displacement will be obtained by adding together the displacements due to the two S.H.M.'s. Thus at a time represented by the point L the total displacement will be equal to $\overline{PL} + \overline{pL}$, while at a time represented by K, the

component displacements being in opposite directions, the total displacement is equal to $\overline{KQ} - \overline{Kq}$, and since $\overline{KQ}$ is equal to $\overline{Kq}$, the displacement is zero. Hence if we construct a curve such that the ordinates are everywhere equal to the algebraic sum of the ordinates of the two component curves, this curve will represent the resultant displacement. The resultant thus obtained is shown dotted in Fig. 42.

In Fig. 43 the same curves are compounded, but the time scale is made smaller, so that more periods of each curve may be shown. It will be seen that the resultant curve, although not a sine curve, is a periodic

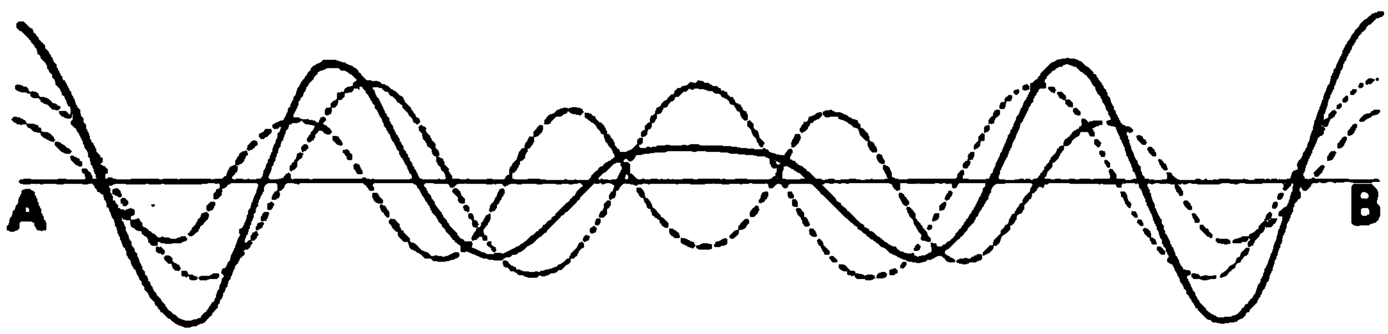


FIG. 43.

curve, and hence the resultant motion is periodic, the period being equal to AB, *i.e.* to five times the period of the quicker vibration, or four times that of the slower.

If the two S.H.M.'s to be compounded are of nearly the same period, say in the ratio of 9 : 10, then the compound harmonic curve obtained will, as shown in Fig. 44, everywhere approximate to the form of a sine curve, but the amplitude will alternately wax and wane; the maxima occurring when the component vibrations are exactly in phase, and the

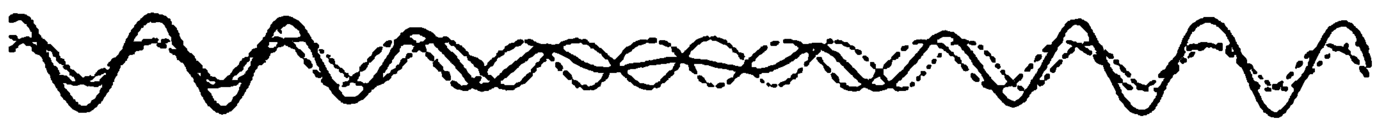


FIG. 44.

minima when the phases differ by half a period. As in 9 periods of the slower vibration there occur 10 periods of the quicker, in this interval one will have gained exactly one period on the other, and they will again be in the same phase. Thus the curve shows that the maxima occur at every 10th period of the quicker vibration.

This waxing and waning of the resultant motion, when two S.H.M.'s of nearly the same period are compounded, is the cause of the phenomenon of beats in music, and will be further studied when we come to the subject of sound.

55. Fourier's Theorem.—In the previous section we have compounded two harmonic curves and drawn a resultant curve. The same method can be employed to compound any number of harmonic curves. The curves having all been drawn, with their appropriate amplitude,

period, and phase, the resultant curve is drawn so that at every point its ordinate is equal to the algebraic sum of the ordinates of the component curves at that point. By suitably choosing the period and amplitude of

FIG. 45.

the component harmonic curves, it is possible, as illustrated in Figs. 45 and 46, to produce a periodic resultant curve of a type very different from a sine curve.

Fourier first showed that any periodic curve, as long as it nowhere goes to an infinite distance from the axis of X, can be built up by compounding together a finite number of harmonic curves the periods of which are commensurate. This last condition is necessary, for otherwise the resultant curve obtained by compounding the curves would never

FIG. 46.

exactly repeat itself, and would not be periodic. Hence it follows that any periodic motion can be considered as the resultant of a number of commensurate S.H.M.'s. If T is the period of the complex periodic motion, then the periods of the component S.H.M.'s will be included in the numbers T, $T/2$, $T/3$, $T/4$, &c.

As an illustration of the way in which a periodic curve of a given form may be built up by the combination of a number of S.H.M.'s, suppose the required curve to be represented by the lines ABCDEFG (Fig. 46). The thick continuous curve given in the figure is obtained by compounding the three S.H.M.'s shown dotted, of which the frequencies are in the ratio 1 : 3 : 5, while the amplitudes are as 1 : 1/3 : 1/5. It will be seen that even with three terms an approximation to the required form is produced. In Fig. 47 the result of combining 100 S.H.M.'s, having frequencies

Fig. 47.

proportional to the numbers 1, 3, 5, 7, 9, &c., and amplitudes proportional to 1, 1/3, 1/5, 1/7, 1/9, &c., is shown on a reduced scale. It will be noticed that in this case the required curve is almost perfectly reproduced.

Machines have been devised, called harmonic analysers, to determine mechanically the amplitudes of the S.H.M.'s of the periods T, $T/2$, $T/3$, &c., required to build up any given curve. Other machines are capable of drawing the resultant of a certain number of S.H.M.'s of given amplitude and period.

PART III—DYNAMICS

CHAPTER VIII

NEWTON'S LAWS OF MOTION

56. Subdivisions of Dynamics.—Up to the present the motion of bodies has been considered quite in the abstract, and although we have assumed that the motion varied in certain ways, we have not inquired into the causes of these variations. We now pass on to consider the effects of force as shown in its action on the motion or equilibrium of material bodies. This branch of the subject of mechanics is called *Dynamics*. Dynamics is sometimes subdivided into two sections; in one, called Kinetics, the effect of forces on the motion of bodies is studied, while in the other, called Statics, the conditions which must exist if a body remains at rest when acted upon by a system of forces are investigated.

57. Stress.—When one portion of matter acts on another portion, so as to influence its state, then the whole phenomenon of the mutual action of the two portions of matter is called in general a *stress*. In certain particular cases the stress has received a special name; thus we have a tension, a pressure, a torsion, an attraction, a repulsion, &c.

The term stress includes the consideration of both the mutually influencing portions of matter; it is, however, sometimes useful to concentrate our attention on one aspect of a stress, namely, the action on one of the portions of matter, so that we regard the stress as something acting on this piece of matter. From this point of view we say that the phenomena which we observe are the effect of *External* or *Impressed Force* on the portion of matter in question, and are due to the ACTION of the other portion of matter. The opposite aspect of the same stress would in this case be called the *reaction* on the other portion of matter. Hence Action and Reaction are simply different aspects of a stress, just as buying and selling are different aspects of one and the same transaction, according as we look at it from the point of view of one or other of the persons taking part in the transaction.

58. Newton's Laws of Motion.—The effect of external or, as it is sometimes called, impressed force on the motion of bodies is defined in three laws which are known as Newton's Laws of Motion. The first of these laws deals with the behaviour of a body when no external force

acts on it. The second tells us how the external force, when acting, may be measured. The third compares the two aspects of a stress, namely, Action and Reaction. These laws are Axioms, and do not admit of *direct* experimental proof; they depend, however, on convictions drawn from experiment, and their truth is universally admitted by those who have sufficient physical knowledge to thoroughly understand their purport.

59. Newton's First Law.—"Every body continues in its state of rest or of uniform motion in a straight line, unless it be compelled by impressed force to change that state."[1]

This law is also known as the law of *Inertia*, since it states that no body is capable of altering its state of rest or of motion without the intervention of some outside influence; and this fact we express in scientific language by saying that every body has inertia.

The law in the first place gives a definition of force, since it states that force is that action by means of which the state of rest or motion of a body is changed, and that unless a force acts no such change will occur. We may therefore define force as that which tends to produce change of motion in a body on which it acts.

In the next place the law tells us how a body will move when it is unacted upon by external forces. It says that if the body is in motion then it will continue moving uniformly in a straight line, if at rest it will continue at rest.

Indirectly the law may be taken as defining equal times. The times which a body, unacted upon by external forces, takes to pass through equal spaces are equal.

Since we are unable to obtain a body which is entirely unacted upon by external force, we cannot experimentally prove that if once set in motion it would continue to move uniformly. We find, however, that the more we reduce the magnitude of the impressed forces acting on a body, the greater is its tendency to continue moving at a uniform rate in a straight line when once it has been set in motion. Thus we know that if a stone is thrown along the surface of a road it will soon lose its motion. If thrown along the surface of smooth ice—in which case the friction, which is an impressed force tending to check the motion, is much less than in the case of the road—it will, however, continue to move very much longer.

A much more powerful argument for the validity of the law is obtained by considering that we can by its means solve problems in mechanics, and the solutions thus obtained always agree with observation, so that we conclude that our fundamental assumption is correct. Thus every one who makes use of the Nautical Almanack to discover the position of a star or the time of an eclipse, tacitly allows the correctness of Newton's law, for it is by the assumption of the correctness of the law that the numbers there given have been calculated.

[1] Corpus omne perseverare in statu suo quiescendi vel movendi uniformiter in directum, nisi quatenus illud a viribus impressis cogitur statum suum mutare.

60. Newton's Second Law.—The first law having stated that it is force alone which can produce change of motion, the second law tells us how the change of motion depends on the magnitude and direction of the force.

Before stating the law in Newton's words, we must consider some definitions which he prefixes to the laws.

(1) The *Quantity of Motion*, or the *Momentum*, of a rigid body moving without rotation is proportional to its mass and its velocity. The reasonableness of this definition will appear if we remember that the effort required to stop a body of great mass, such as a railway train, when moving with given velocity, is much greater than that required to stop a body of small mass, say a marble, when moving with the same velocity. Again, a greater effort is required to stop a bullet projected from a rifle with a high velocity than to stop a similar bullet when simply thrown by hand, and thus moving with a comparatively slow velocity. If, then, we take as the unit of momentum that of unit mass moving with unit velocity, the momentum of a mass m moving with a velocity v will be mv. The dimensions of momentum are $[L^1M^1T^{-1}]$.

The change in momentum of a body is proportional to the mass of the body and the change in velocity. This follows at once, since the mass of body cannot alter; hence the only thing that can effect the magnitude of the momentum is a change in velocity. The *rate* of change of momentum is proportional to the mass and the acceleration (since the acceleration is the rate of change of the velocity). It must be remembered that the term velocity is used in the above in its most general sense (§ 30), and hence the momentum of a body changes when the direction of motion changes, although the speed may remain constant.

We may now state Newton's second law :—"Change of motion is proportional to the impressed force, and takes place in the direction of the straight line in which the force acts."[1] By motion Newton means quantity of motion or, as it is now called, momentum, and in the same way the term impressed force includes the idea of time, for the magnitude of the change of momentum produced will depend on the time during which the force acts as well as on the magnitude of the force. The product of the magnitude of a force into its time of action is called the *impulse* of the force. Hence we may restate the first part of the law and say : Change of momentum is proportional to the impulse of the impressed force. It is important to notice that this law states that it is the *change* in momentum which is proportional to the impulse of the force, and hence it is immaterial whether the body on which the force acts is originally at rest or in motion in any direction ; the *change* in its momentum in the direction in which the force acts is always proportional to the impulse of the force.

[1] Mutationem motus proportionalem esse vi motrici impressae, et fieri secundum lineam rectam qua vis illa imprimitur.

Thus, suppose we have a body of mass m at rest, and that a force acting in the direction from South to North imparts a velocity v to the body in unit time, so that the momentum generated is mv. If, instead of being at rest, the body had been moving with a velocity v from South to North when the force commenced to act, then at the end of a second it would be moving with a velocity $2v$ from South to North, since in this time its momentum must have changed as much as it did before, and the change in momentum is equal to the product of the mass into the change in velocity, and $(2v-v)m$ is equal to mv. If the body were originally moving with a velocity v from North to South, *i.e.* in an opposite direction to that of the force's line of action, then at the end of a second the body would be at rest, having lost mv units of momentum in a direction opposite to the line of action of the force, which is equivalent to the gain of mv units in the direction of the line of action. If the body were originally moving with a velocity v from East to West, then at the end of a second it would have gained mv units of momentum in the South to North direction, that is, since it originally had *no* momentum in this direction, its original velocity having no component in this direction, it will now have a component velocity in the South to North direction of v units. Further, it is immaterial whether the body is under the influence of *one* force or several. However many the forces acting on the body, each force will produce the same change of momentum in its own direction that it would produce supposing it alone acted.

The second law gives us a means of defining and measuring forces as well as masses. If we have a number of forces, then, according to Newton's second law, the changes of momentum which they would separately produce in a given time are proportional to the forces. Hence if they all act in succession on the *same* mass, the changes in the velocity produced will be proportional to the forces, so that we may measure the relative magnitudes of the forces by determining the change in velocity each force will produce in a given mass in a given time.

On the other hand, if we allow a given force to act in succession on a number of different masses for a given time, then, since it will in each case produce the same change in momentum, the value of the product of the mass of each body into the change of velocity produced by the force is the same for all. The changes in velocity produced are therefore inversely as the masses of the bodies in which these changes are produced by the same force acting for the same time.

61. Unit of Force.—Since a force is measured by the change in momentum it produces in its line of action, if a force when acting on a mass m for a time t changes the velocity in its line of action by v units of velocity, then the force is measured by the quotient mv/t. It will have been observed that, in the statement of the law and in the remarks that have been made, it is said that the force is *proportional* to the change in momentum produced in unit time. It is, however, very convenient to so

choose the unit of force that the value of a force is numerically equal to the change in momentum produced in a second. The unit force will then be such that it produces in unit mass unit changes of velocity per second, *i.e.* unit acceleration. Hence if a force F acting on a mass m for a time t changes the velocity in its own direction from v_1 to v_2, we have

$$F=\frac{m(v_2-v_1)}{t}.$$

or, since $(v_2-v_1)/t$ is the acceleration (a) produced,

$$F=ma.$$

The dimensions of force can be obtained from this equation by introducing the symbols for the units and then making F, m, and a each unity. Thus

$$\begin{aligned} F[F]&=m[M]a[LT^{-2}] \\ [F]&=[M][LT^{-2}] \\ &=[L.M.T^{-2}]. \end{aligned}$$

In the *c.g.s.* system the unit force is such that it produces an acceleration of one centimetre per second per second in a mass of one gram, and is called a *dyne.*

In the foot-pound-second absolute system of units the unit of force is such that it produces an acceleration of one foot per second per second in a mass of one pound, and is called a *poundal.*

We may make use of the dimensions of a force to determine the relation between the dyne and the poundal. Suppose a given force to be equal to d dynes or p poundals, and further that L, M, and T are the units of length, mass, and time in the *c.g.s.* system, and L_1, M_1, T_1 those in the foot-pound-second system. Then, since the actual magnitude of the force must be the same whatever the units used to measure it, we have

$$d[LMT^{-2}]=p[L_1M_1T_1^{-2}]$$

$$\therefore \quad \frac{d}{p}=\left[\frac{L_1}{L}\right]\cdot\left[\frac{M_1}{M}\right]\cdot\left[\frac{T^2}{T_1^2}\right]$$

Here $\frac{L_1}{L}$ is the ratio of a foot to a centimetre, and is equal to 30.48 (§ 11), while $\frac{M_1}{M}$ is the ratio of a pound to a gram, and is equal to 453.59 (§ 12), while $\frac{T^2}{T_1^2}$ is unity, since the second is the unit of time in either case. Substituting these values, we get

$$\begin{aligned} \frac{d}{p}&=30.48\times 453.593 \\ &=13825.5. \end{aligned}$$

Hence the number of dynes in the given force is to the number of poundals as 13825.5 : 1. If therefore the given force is one poundal,

the number of dynes it contains is 13825.5, so that one poundal = 13825.5 dynes.

62. Impulsive Force.—In certain cases the force acts for so short a time that we are unable either to measure its magnitude or the time during which it acts. It is, however, in these cases generally possible to measure the total effect of the force in changing the motion of the body on which it acts. Now, as has been stated in § 60, the total effect of a force in changing the motion of a body, or the impulse of the force, is measured by the change in momentum produced. Hence in the case of these forces of very short duration the impulse will be used to measure the effect of the force; and this is equal, if the force is uniform, to the product of the force into its time of action, or, if the force is variable, to the product of the mean value of the force into the time of action. Thus forces of short duration, as for example that exerted by a blow of a hammer, were originally called impulsive forces; and it was in this connection that the term impulse was originally used. There is, however, no essential difference between such a force and forces which last for a longer interval, the only distinction being that in the one case, from lack of experimental means, we are unable to make the necessary measurements. The term impulse is, therefore, now used in the more general sense, as applicable to the product of any force into its time of action.

63. Newton's Third Law.—"To every action there is always an equal and contrary reaction: or, the mutual actions of any two bodies are always equal and oppositely directed."[1] In this law the word action is used to represent the one aspect of a stress spoken of in § 57, and the word reaction to represent the other. Hence Newton's third law states that all forces are of the nature of a stress between portions of matter, since it states that every force must necessarily be accompanied by an equal and oppositely directed reaction.

If you press your finger on the table, you feel the table pressing your finger. In the case of a horse towing a boat, the forward pull exerted by the horse on the tow-rope is exactly equal to the backward pull exerted by the tow-rope on the horse. Many people find a difficulty in accepting the above statement with reference to the equality of the action and reaction in the case of a horse towing a boat, since they think that if the force exerted by the horse on the rope were not a *little* greater than the backward force exerted by the rope on the horse, the boat would not progress. In this case we must, however, remember that, as far as their *relative positions* are concerned, the horse and the boat are *at rest*, and form a single body, and the *action* and *reaction* between them, due to the tension on the rope, must be equal and opposite, for otherwise there would be relative motion, one with respect to the other. The horse obtains the necessary purchase to move both itself and the boat where

[1] Actioni contrariam semper et aequalam esse reactionem: sive corporum duorum actiones in se mutuo semper esse aequales et in partes contrarias dirigi.

its feet touch the ground. At these points the horse's hoofs exert a force which has a component in a backward direction, the corresponding reaction of the ground having a component in the forward direction; and it is this component which produces the motion of the horse and boat.

64. Action at a Distance.—Since force is always part of a stress, and is only produced by the agency of one portion of matter on another, it is of interest in every case to examine the mechanism by means of which this influence of one piece of matter on another is carried on. In some cases, such as that where two portions of matter are connected by a stretched string, it is quite evident by what means the one piece of matter exerts a force on the other, for it is by the stretched string. In other cases, however, with which we shall deal more fully later on, one piece of matter acts on another, and is reacted on by that portion of matter, but without our being able to detect any intermediate body which plays the part of the string in the first example. As an instance, we may take the case of the force exerted by a magnet upon a piece of iron, even when they are at some distance apart. In this case the force still exists if we remove, as well as we are able, all matter that can be detected by our senses, and which for short, and for the reasons given later, may be called ponderable matter, from the space between the magnet and the iron, or if we place other portions of matter between the two.

It was at one time considered sufficient in a case such as the above to say that the magnet exerted a force on, or acted on, the iron at a distance, and to dismiss the question by saying that this was a case of magnetic attraction.

If a conjurer makes a portion of matter, say a block of wood, follow his hand about, we at once say that he has a string or some other mechanism connecting the block to his hand, and although we are quite unable to see the nature of this connecting link, we may be satisfied in our own mind that it does really exist. In the same way, since we are unable to think of one portion of matter acting upon another portion of matter without something connecting the two, by means of which the action is transmitted, it is natural to suppose that there exists some mechanism, or, as it is called, a *medium*, by means of which the action of the magnet on the iron is transmitted. In this way of looking at the subject we suppose that the action is transmitted by each portion of the medium affecting that which lies next to it, and so handing on the action till the second portion of matter is reached. Here, then, we only assume action in proximity. Since, however, no one has yet been able to imagine the constitution of a medium which shall be capable of transmitting all the different kinds of action which experiment shows matter to produce on matter, we ordinarily use the language of the theory which supposes that matter can act upon matter at a distance, without any connecting mechanism, by means of an agent which we call force. Although no one has been able to imagine the necessary medium, nevertheless we firmly

believe that such a medium does exist; and a considerable portion of the present volume will consist of a description of experiments which have been made with a view of determining the properties of this medium.

Newton made an experiment to show that in the case of the action exerted by one portion of matter on another at a distance the third law of motion was true. He floated a magnet and a piece of iron on water by placing them on two portions of cork, so that these pieces of cork were in contact. He found that neither the magnet nor the iron was able to move the other along. Hence the magnet must be attracted by the iron just as much as it attracts the iron.

65. Graphical Representation of a Force.—In order to completely define a force, we require to know three things about it: (1) Its point of application; (2) its direction; and (3) its magnitude.

All these three particulars can be represented by a straight line, for we may draw a straight line through the point where the force acts in the direction in which the force tends to cause the momentum of the body to increase, and so that this line contains as many units of length as there are units of force in the force. In order to indicate in which sense along the line the force acts, it is usual to place an arrow-head with the point turned in the way of action of the force.

When we represented a velocity by a line, it was mentioned (§ 36) that all equal parallel lines represented the same velocity. In the case of forces, since the line has to be drawn through a definite point, the point of application, we can only draw a single line to represent any given force. It is however sometimes convenient, when we have a number of forces acting, to draw lines to represent the magnitude and direction of the forces only, so that all equal and parallel lines represent the same force. In such a case we must be very careful to remember that these lines do not *completely* represent the forces.

66. Composition of Forces Acting at a Point.—A force being a vector quantity, we compound two forces which act at a point by the parallelogram method. This also follows from Newton's second law. For suppose a_1 and a_2 are the accelerations, the two forces F_1 and F_2 would respectively produce in their own direction when they act on a mass m, so that by the second law $F_1=ma_1$ and $F_2=ma_2$. The mass m is therefore moving with an acceleration a_1 in the direction of F_1, and with an acceleration a_2 in the direction of F_2. From O (Fig. 48) draw $\overrightarrow{OP}$ to represent a_1 in magnitude and direction, and $\overrightarrow{OQ}$ to represent a_2 in magnitude and direction. Then the resultant acceleration will be represented by the diagonal $\overrightarrow{OR}$ of the parallelogram constructed on $\overline{OP}$ and OQ. Now the acceleration represented by $\overrightarrow{OR}$, say a, would be produced by a

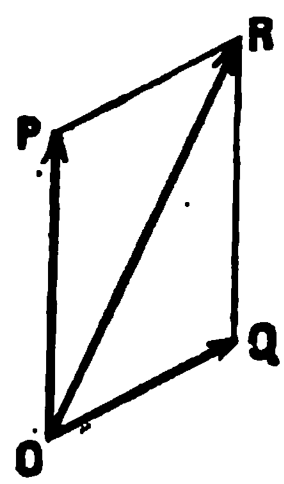


FIG. 48.

force ma, acting in the direction $\overrightarrow{OR}$, and hence this force is the resultant of F_1 and F_2.

Since
$$\overrightarrow{OP}:\overrightarrow{OQ}:\overrightarrow{OR}:: \ a_1: \ a_2:a$$
$$\therefore \quad \overrightarrow{OP}:\overrightarrow{OQ}:\overrightarrow{OR}::ma_1:ma_2:ma$$
$$\therefore \quad \overrightarrow{OP}:\overrightarrow{OQ}:\overrightarrow{OR}:: \ F_1: \ F_2:R.$$

Hence, to find the resultant of two forces acting at a point we need not consider the accelerations they would produce, but if we draw from a given point two straight lines to represent the two given *forces* and complete the parallelogram, the diagonal through the given point will represent the resultant force in magnitude and direction. Since the two component forces acted at one point, the resultant will act at the same point.

If there are any number of forces acting at a point, we can find their resultant by a method similar to the polygon of velocities or accelerations. For if in succession we draw straight lines to represent in magnitude and direction each of the forces, starting from some given point, and draw the line representing each subsequent force from the point where the line representing the previous force ended, then the straight line joining the starting-point to the end of the last line so drawn will represent the resultant of all the forces both in magnitude and direction. As before, the point of application of the resultant must be at the same point as that of the component forces.

67. Resolution of Forces.—A force, like a velocity or an acceleration, can be resolved into components along any given directions, the method employed being exactly the same as that given in § 38 for the case of velocities.

The usual case is to resolve a force into components along two directions at right angles to one another. As an example of the resolution of forces, we may take the case of a boat sailing in any direction except directly before the wind. Let BR (Fig. 49) represent a boat, and AD the plan of the sail. If WAC is the direction of the wind, and we take AC to represent the force the wind would exert on the sail if it were placed at right angles to the direction of the wind, we must resolve this into components, one ($\overrightarrow{AE}$) perpendicular to the sail, which is alone efficacious as far as the action of the wind on the sail is concerned, and the other parallel to the sail AD.

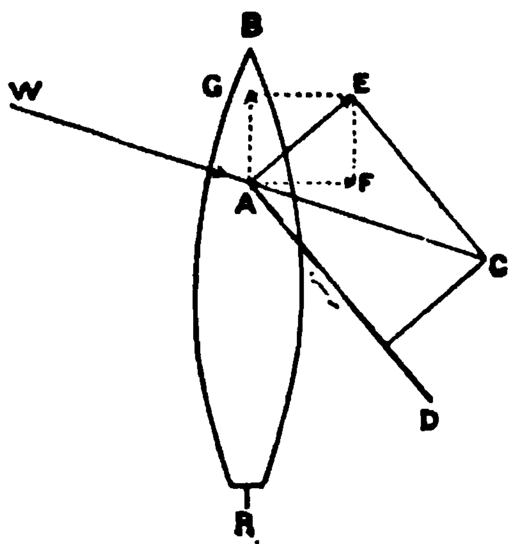


FIG. 49.

The force $\overrightarrow{AE}$ has now to be resolved along and across the boat. The component $\overrightarrow{AG}$ is alone effective as far as the headway is concerned.

The component $\overrightarrow{\text{AF}}$, at right angles to the course of the boat, tends to make the boat travel through the water in a direction at right angles to its length, *i.e.* it produces leeway.

68. Moment of a Force.—When a force acts on an extended body it produces, in general, both motion and deformation of the body, *i.e.* strain. As we are not at present dealing with the question of strain, we shall consider the body to be rigid (§ 43). In the case of an extended rigid body a force will, in general, produce both a motion of translation and a motion of rotation. If the force is so applied that translation only takes place, the body moves as if it were a particle having a mass equal to that of the body concentrated at a certain point, which is called the centre of inertia or centre of gravity. Hence if any number of forces act on a rigid body so that their resultant passes through the centre of inertia of the body, the motion they will generate will be a pure translation. If the direction of the resultant does not pass through the centre of inertia, there will be a motion of rotation produced as well as one of translation. In order to simplify matters when studying the effect of a force in producing rotation, we shall suppose that all motion of translation is prevented by having a point or sometimes a line in the body kept fixed (§ 47).

The effect of a force F in producing rotation depends not only on the impulse of the force $(F \cdot t)$, but also on the distance between the line of action of the force and the axis about which rotation is capable of taking place. Thus it requires a much smaller force to close a door if the force is applied at right angles to the door and near the handle, than if applied near the hinge. It will be noticed that it is the perpendicular distance of the *direction* of the line of action of a force from the axis of rotation, and not the distance from the point of application of the force to the axis, which settles the amount of the turning power of a force. Thus let AB (Fig. 50) be a rigid body capable of rotating about an axis through A perpendicular to the paper. Then the turning power of a force acting along F_1 and applied at B is much less than that of an equal force acting along F_2, although the distance between the point of application and the axis is the same in each case. The turning power or torque depends on the magnitude of the force and the perpendicular distance between the axis and the direction of the force, *i.e.*, on $\overline{\text{AC}}$ or $\overline{\text{AD}}$ in the above figure.

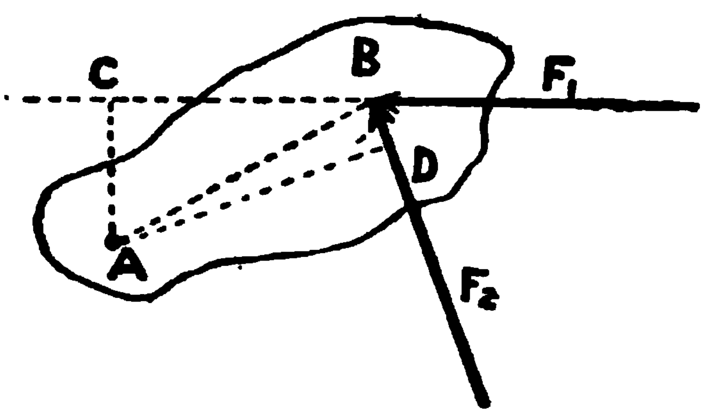


FIG. 50.

The product of the magnitude of a force into the perpendicular distance between the axis and the direction of the force is called the *moment* of the force. Hence in the above example the moments of the forces are $F_1 . \overline{\text{AC}}$ and $F_2 . \overline{\text{AD}}$ respectively. The moment of a force is positive if it

tends to produce rotation in the positive (*i.e.* anti-clockwise) direction, and negative if it tends to produce rotation in the negative direction.

Since the resultant of any number of forces is by definition,[1] a single force which is capable of replacing the component forces as regards their effect, it follows that the moment of the resultant about *any* point must be equal to the sum of the moments of the components about the same point, or otherwise the resultant would not correctly replace the turning moment of the components.

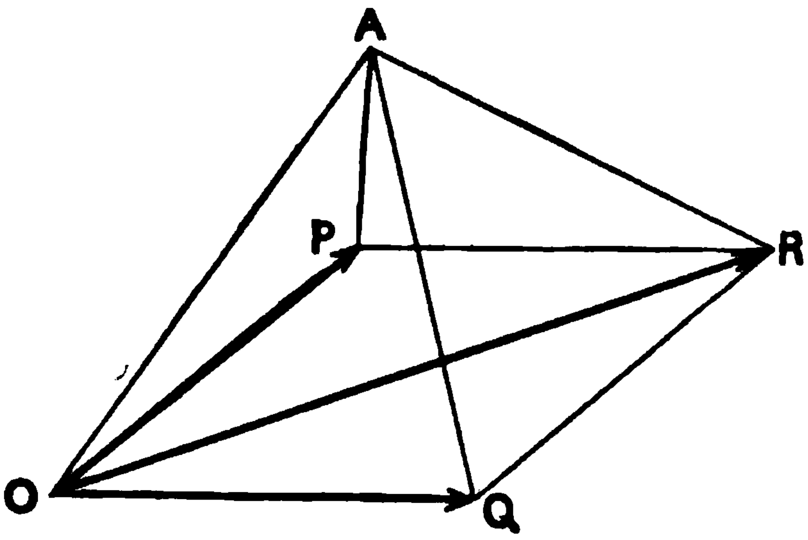


FIG. 51.

That the parallelogram construction, and hence also the polygon of forces, gives a resultant fulfilling the above condition may easily be shown in the following way. Let $\overrightarrow{OP}$, $\overrightarrow{OQ}$ (Fig. 51) be two forces acting at O and $\overrightarrow{OR}$, the resultant obtained by completing the parallelogram; and let A be any point in the plane about which moments are to be taken. Join AO, AP, AQ, and AR. Then the triangle AOQ is equal to the sum of the triangles APR, OPR, for they are on equal bases $\overline{OQ}$, $\overline{PR}$; and the height of the one triangle is equal to the sum of the heights of the other two. This may be proved thus: if b is the length of the base of each triangle, and h_1 and h_2 the heights of the triangles APR, OPR respectively, then $\frac{1}{2}b(h_1+h_2)=\frac{1}{2}bh_1+\frac{1}{2}bh_2$, that is, the area of the triangle on base b and height (h_1+h_2) is equal to the sum of the areas of the triangle on base b, having heights h_1 and h_2 respectively. The triangle AOR is obviously equal to the sum of the triangles AOP, APR, OPR. Hence the triangle AOR is equal to the sum of the triangles AOP and AOQ. Now the area of the triangle AOR is equal to half the product of the base $\overline{OR}$ into the perpendicular distance between A and OR. The product of $\overline{OR}$, which represents in magnitude and direction the resultant of the forces, into the perpendicular distance between A and OR, is the moment of the resultant about A. Hence the moment of $\overrightarrow{OR}$ about A is represented by twice the area of the triangle AOR. In the same way the moments of $\overrightarrow{OP}$ and $\overrightarrow{OQ}$ about A are represented by twice the areas of the triangles AOP and AOQ. Hence it follows, from the relation between the areas of these triangles found above, that the moment of $\overrightarrow{OR}$ about A is equal to the sum of the moments of $\overrightarrow{OP}$ and $\overrightarrow{OQ}$ about A.

[1] In § 70 the conditions that two forces cannot have a single resultant will be discussed.

69. Composition of Parallel Forces.—In the case of an extended body acted upon by two forces whose directions are parallel, the resultant force will be equal to the algebraic sum of the two components, *i.e.* to the arithmetical sum if they act in the same sense, and to the difference if they act in opposite senses. As far as the motion of translation of the body is concerned, this is all that is required. If, however, we require the effect of the forces in producing rotation, it is further necessary to know the point of application of the resultant.

In order to find the position of the resultant, we make use of the fact that the moment of the resultant about any point must be equal to the sum of the moments of the components about the same point. If the point chosen is on the line of action of the resultant, then, since in this case the moment of the resultant about this point is zero, the moments of the components about this point must be equal and opposite. If the parallel force act in the same sense, the resultant will lie between them, for in this case the moments of the forces about any point between them will be of opposite sign, since they will tend to cause rotation in opposite directions. If the forces act in opposite senses, then the resultant must lie outside the forces, and on the side of the larger force, for then the distance between the larger force and any point on the resultant will be *less* than that between the smaller force and this point, so that the smaller force acting at a greater distance may have a moment equal to the larger force acting at a smaller distance.

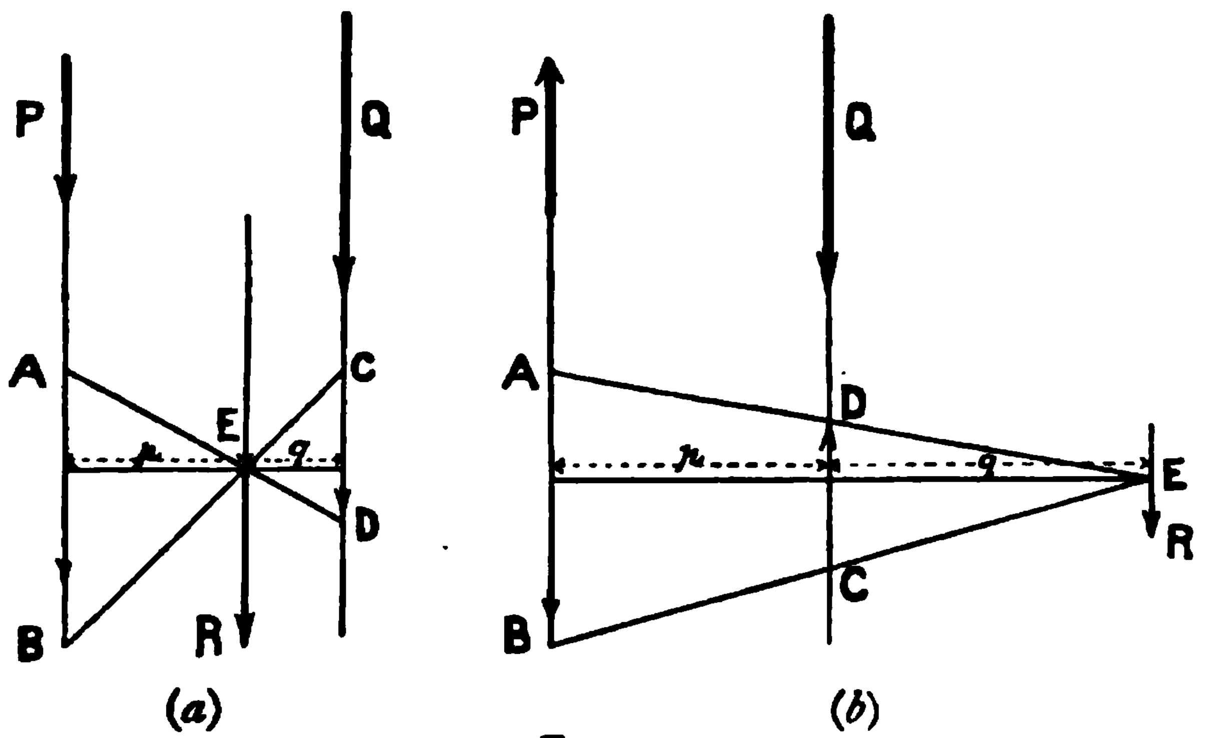


FIG. 52.

The most generally useful construction for finding the position of the resultant is as follows: Let P and Q (Fig. 52) represent the forces. Anywhere along the line of action of P cut off a portion $\overline{AB}$ equal in length to Q, and somewhere in the line of action of Q cut off $\overline{CD}$, equal in length to P. Then join AD and BC. If the forces act in the same sense, join

crosswise, as at (*a*), so that the point of intersection E lies between the forces; if they are in opposite senses, join them without crossing, as at (*b*). Then in either case the resultant will pass through the point E.

The triangles ABE and DCE are similar, so that their heights are to one another as their bases. Hence if p is the height of the triangle ABE, and q that of the triangle DCE, we have

$$\frac{p}{q}=\frac{\overline{AB}}{\overline{CD}}=\frac{Q}{P}.$$

$$\therefore \qquad Pp=Qq.$$

But p is the perpendicular distance between the point E and the force P. Therefore Pp is the moment of P about E. In the same way Qq is the moment of Q about E. And we see from the above that the moments are equal in magnitude; that they are of opposite sign is obvious from the figures. Hence the resultant is parallel to the forces, and passes through E.

70. Couples.—If the two parallel forces are equal, the resultant must be at an equal distance from each, so that when the forces are in opposite senses it must be at an infinite distance, for otherwise, as it has to be outside the two forces, it would be nearer one than the other, so that the moments would not be equal. The magnitude of the resultant, however, is in this case zero. As far as translation is concerned, a system consisting of a pair of equal and opposite parallel forces, which is called a *couple*, will produce no result; it may, however, produce rotation.

Let P and P′ (Fig. 53) represent the equal and opposite parallel forces, and from E, any point in their plane, draw EAB perpendicular to the direction of the two forces meeting them at A and B. Then the moment of P about E is $-P.\overline{BE}$ (for sign see § 68), and that of P′ is $+P'.\overline{AE}$ in (*a*), and $-P'.\overline{AE}$ in (*b*). Hence the sum of the two moments in (*a*) is $-P.\overline{BE}+P'.\overline{AE}$, which, since $P'=P$, is equal to $-P.\overline{AB}$. In (*b*) the sum of the moments is equal to $-P.\overline{BE}-P'.\overline{AE}$, which is equal to $-P.\overline{AB}$. Hence we see that the sum of the moments about any point is constant, and is equal to the product of one of the forces into the distance between the lines of action of the forces. This product is called the moment of the couple.

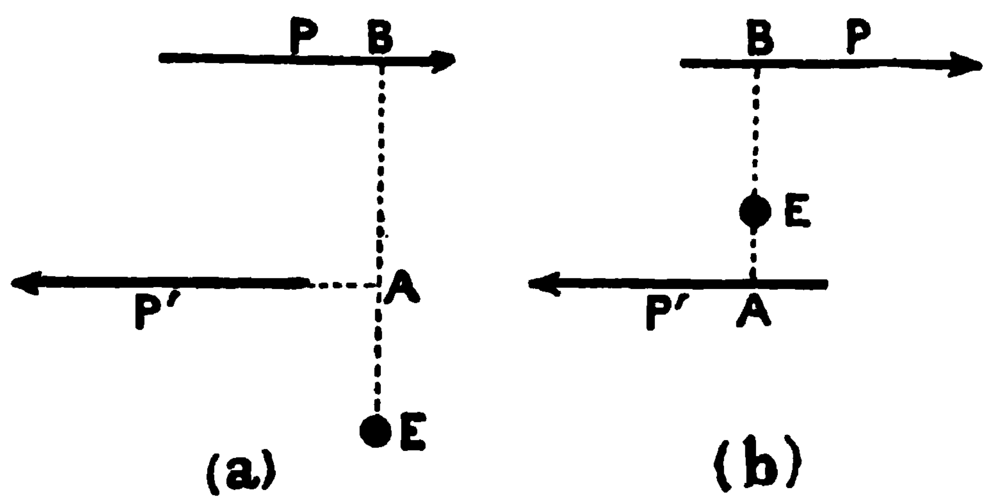


FIG. 53.

We are now in a position to generalise, and say that any number of

forces acting on a rigid body may be replaced by a single force acting through the centre of inertia of the body, and which is alone effective in producing motion of translation, and a couple which is alone effective in producing motion of rotation. For, taking any two of the forces which are not equal and opposite parallel forces, we may replace them by their resultant. This resultant can then be combined with one of the remaining forces, and so on till finally we have left either (1) a single force passing through the centre of inertia, when translation only takes place; or (2) two equal and opposite parallel forces, which produce rotation only; or (3) a single force which does not pass through the centre of inertia. In this last case, if we add two equal and opposite forces acting through the centre of inertia and parallel to the resultant, they will not influence the motion. One of these forces will then form a couple with the resultant, and the other will be a force equal and parallel to the resultant, acting through the centre of inertia and tending to produce translation of the body.

CHAPTER IX

EQUILIBRIUM OF FORCES

71. Equilibrium.—When the forces which act on a body are so balanced that they produce no acceleration in the body, that is, do not alter its state of motion, they are said to be in equilibrium. A study of the conditions that have to be fulfilled in order that the forces considered may be in equilibrium is sometimes considered as a separate branch of mechanics, called Statics.

The name statics is at first sight rather misleading, since it does not follow because the forces acting on a body are in equilibrium that the body is at rest, for if the body is originally moving the velocity will continue uniform, and not be altered by the forces. The appropriateness of the name, however, is apparent, if we consider that unless the forces acting on a body are in equilibrium it is impossible for the body to *remain* at rest. Hence we may if we like define forces in equilibrium as such that they render it possible for the body on which they act to remain at rest.

72. Conditions for Equilibrium of a Particle.—It is obvious that a particle acted upon by a single force cannot be in equilibrium.

For two forces acting on a particle to be in equilibrium, they must fulfil the following conditions: They must be (1) equal in magnitude, (2) act along the *same* straight line, (3) be of opposite sense. When referring to these conditions in future, we shall for shortness simply say that the forces must be equal and opposite, but it must be remembered that this is only an abbreviation for the above three conditions.

The condition that three forces acting on a particle may be in equilibrium is that any one of the forces must be equal and opposite to the resultant of the remaining two, for we may, if we please, replace any two of the forces by their resultant, when we should have reduced the problem to the equilibrium of two forces. The resultant of any two of the forces, say P and Q, must lie in the plane containing P and Q. Hence if there is to be equilibrium the third force, since it must be equal and opposite to this resultant, must also lie in the plane containing the other two forces. Hence the first condition for equilibrium is that the three forces must all lie in one plane. As to the relations between the magnitude of the forces, the resultant of any two (P and Q) is represented by the diagonal $\overrightarrow{OR}$ (Fig. 54) of the parallelogram constructed

on the lines $\overline{OP}$ and $\overline{OQ}$ as adjacent sides. Hence the other force must be represented in magnitude and direction by $\overrightarrow{RO}$, or by $\overrightarrow{OS}$, where $\overline{OS}$ is equal to $\overline{OR}$ and in the same straight line with it. Since $\overline{QR}$ is equal to $\overline{OP}$, we may take $\overrightarrow{QR}$ to represent the force P in magnitude and direction (§ 65). Then the three forces will be represented by $\overrightarrow{OQ}$, $\overrightarrow{QR}$, and $\overrightarrow{RO}$, the sides of a triangle. Hence if it is possible to draw a triangle

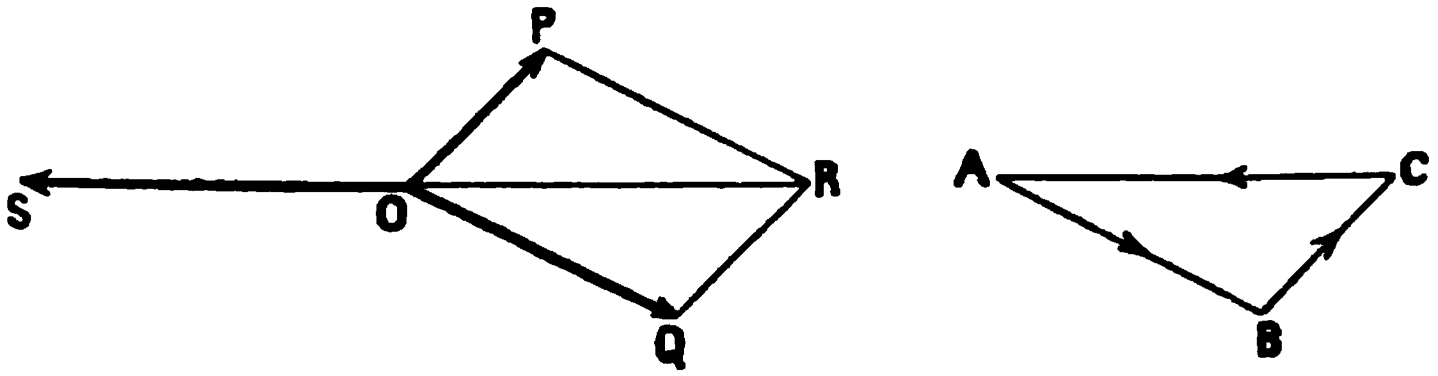


FIG. 54.

of which the sides are parallel or perpendicular to the three forces and proportional to them in magnitude, the forces will be in equilibrium. It must be specially noticed that in drawing the triangle the sides must all be drawn in the same sense as the forces, so that when we place arrows on the sides to show in which sense the forces act, all the arrows may point the same way *round* the triangle, as shown at A B C in Fig. 54.

The conditions of equilibrium for any number of forces acting on a particle are that the forces can be represented in magnitude and direction by the sides of a closed polygon taken in order, *i.e.* drawn in the same sense as the forces. This at once follows from the polygon of forces, for the resultant of all the forces but one is represented in magnitude and direction by the line joining the starting-point to the end of the last line drawn in the polygon, *i.e.* by the remaining side of the polygon, which by supposition represents in magnitude the only force not yet included, but is in an opposite sense.

73. Conditions for Equilibrium of a Rigid Body.—In the case of a rigid body the line of action of the forces need not all pass through a single point, and in order that the body may be in equilibrium the forces must not tend to produce either translation or rotation. If the directions of all the forces pass through a single point they cannot produce rotation, and hence if they fulfil the conditions given in the preceding section for a particle they will be in equilibrium. If, however, the lines of action of the forces do not all pass through a point, then, in order that there may be no rotation, they must have no resultant moment tending to turn the body about any axis. The general condition for equilibrium is therefore that the sum of the moments of all the forces

taken about *every* point must be zero, and that the forces can be represented in magnitude and direction by the sides of a closed polygon taken in order.

Since in most cases we shall only have to deal with forces acting in a plane, it is of interest to examine the condition for equilibrium in this case a little more fully. As by supposition the forces all act in a plane, it is evident that they can only tend to produce motion in this plane (by Newton's second law). Hence if we take two fixed lines not parallel (and preferably at right angles) in this plane, every possible translation must either be parallel to one or other of these lines, or else compounded of translations parallel to the two. Hence if the sum of the components of the forces when resolved parallel to these lines is zero, there will be no tendency to motion along either of these directions, so that there will be no translation. The condition for no rotation is that the sum of the moments about every point in the plane shall be zero. If both conditions are fulfilled there is equilibrium. If only the first condition holds, then there is rotation without translation, *i.e.* all the points of the body move in circles about a fixed point as centre; if the second condition alone is fulfilled, then there is translation without rotation, *i.e.* all the points of the body move with the same velocity in parallel paths.

CHAPTER X

WORK AND ENERGY

74. Definition of Work.—When a force acts upon a body, and the point of application of the force moves in the direction of the line of action of the force, the force is said to do *work on* the body. The amount of work done by the force is measured by the product of the force into the distance, measured along its line of action, moved through by its point of application. Hence if a force F acts on a body while its point of application moves through a distance s in the line of action of the force, the work (W) done *by* the force is given by

$$W = Fs.$$

If the body moves through a distance s in the direction opposed to the force, work is said to be done *against* the force, the work done being as before measured by the product Fs.

If the displacement of the point of application of the force is not along the line of direction of the force, but inclined to it, then we must calculate the component of the displacement along the direction of the force, and this component multiplied by the force gives the work done either by or against the force, as the case may be, during the displacement. Thus suppose $\overrightarrow{AC}$ (Fig. 55) represents the direction of the line of action of the force (F) and $\overrightarrow{AB}$ the displacement of the point of application. Then the component of the displacement along the line of action of the force is $\overrightarrow{AD}$, obtained by drawing BD perpendicular to AC. Hence the work done is $F.\overline{AD}$. The correctness of the above construction is evident, for the dis-

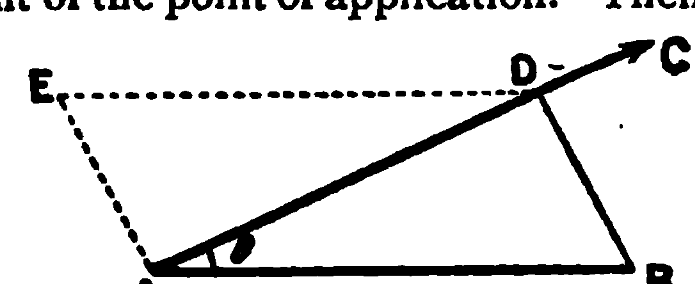


FIG. 55.

placement $\overrightarrow{AB}$ of the point of application may be replaced by the displacements $\overrightarrow{AD}$ and $\overrightarrow{DB}$. During these displacements, no work will be done by the force while the point of application is moving from D to B, since the movement is at right angles to the line of action of the force. If we call the angle between the line of action of the force and the line of

displacement of the point of application θ, then $\overline{AD} = \overline{AB} \cos \theta$. Hence the work ($W$) is given by the equation

$$W = \overline{AB}. \cos \theta. F.$$

If the force F were resolved along AB and perpendicular to AB, *i.e.* along AE, the component along AB is to F as $\overline{AB}$ is to AD. Therefore the component along $AB = F. \overline{AD}/\overline{AB} = F. \cos \theta$. Hence we see that $F \cos \theta$ is the component of the force along the direction of the displacement of the point of application, and the work done is equal to the product of this component into the actual displacement of the point of application.

Since $W = Fs$, we have $F = W/s$, or the force acting on a body moving in a straight line is the work done during a given displacement divided by that displacement. Hence we might define force as the *space* rate at which work is done (see note, p. 27).

75. Units of Work.—In the *c.g.s.* system, the unit of work is done when a body acted upon by a force of one dyne moves through a centimetre in the direction of the force. This unit of work is called the *erg*.

As will be shown later on, the dyne is very nearly equal to the force with which the earth attracts (*i.e.* the weight of) one milligram ($\frac{1}{1000}$ grm.), so that approximately an erg of work is done when a milligram is raised through one centimetre. Since an English penny piece has a mass of about 9450 milligrams, and the average height of a table is about 72 centimetres, it follows that about $9450 \times 72 = 680{,}400$ ergs are done when a penny is raised from the floor to the top of a table. It will be seen that the erg is an excessively small unit, and hence for most practical purposes it is usual to use as the unit of work the joule, which is equal to 10,000,000 or 10^7 ergs.

In the foot-pound-second system, the unit of work is done when the point of application of one poundal is moved, in the direction of the force, through a distance of one foot. This unit is called the foot-poundal.

The dimensions of work are:

$$[Force]\,[L] = [L^2.M.T.^{-2}].$$

A foot-poundal is equal to 4.214×10^5 ergs. This equivalent could be at once obtained from the ratio of the pound to the gram, and of the foot to the centimetre, by aid of the dimensions, as was done in § 61 in the case of the absolute units of force.

76. Gravitational Units.—It will be convenient to anticipate in some measure a few of the points which will be dealt with at greater length in Chapter XIII.

It is a matter of common observation that *all* matter is attracted by the earth, or in other words possesses weight. It can be proved by experiment, as we shall see later, that all bodies when allowed to fall freely, that is to move under the influence of the force exerted on them

owing to the attraction of the earth, move with the same acceleration at any given point on the earth's surface. This acceleration of a freely falling body is generally indicated by the symbol g. Let W be the weight of a body of mass m, that is the force with which the earth attracts it, measured in absolute units (that is in dynes if m is measured in grams, and in poundals if m is measured in pounds). Then since the force W when acting on the mass m produces an acceleration g, we have

$$W = mg.$$

The value of g in the *c.g.s.* system is about 981 cm./sec^2., so that the weight of a gram is equal to about 981 dynes. In the foot-pound-second system g is about 32 feet/sec^2., so that the weight of a pound is about 32 poundals.

A system of units is often employed in which the unit of force is taken as the force with which the earth attracts a given lump of matter when it is at a certain fixed point on the earth's surface. This unit of force is then taken as one of the fundamental units, the others being those of length and time. Such a system of units is called a gravitational system, and it is this system which is almost exclusively used by engineers.

The unit of force in the metric gravitational system is the force with which the earth attracts a mass of a gram when at the sea-level and at latitude 45°.[1] This force is equal to a force of 980.6 dynes, for the value of g at sea-level and latitude 45° is 980.6 cm./sec^2.

The unit of force in the British gravitational system is the force with which the earth attracts a mass of a pound at sea-level and latitude 45°. This unit of force is equal to 32.172 poundals, for g has the value 32.172 foot/sec^2.

In the gravitational system of units, since the unit of force is taken as one of the fundamental units in place of mass, the unit of mass is derived from this unit by means of Newton's second law. Thus the unit of mass on the gravitational system when acted upon by the unit force in this system must move with unit acceleration, that is, one centimetre per second per second, or one foot per second per second, as the case may be. Now at sea-level and latitude 45° the quantity of matter (gram or pound) used to define the gravitational unit of force, if allowed to fall freely, would move with an acceleration g (980.6 cm./sec^2. or 32.172 foot/sec^2.), and under these circumstances it would be acted upon by the gravitational unit of force. Hence to move with unit acceleration the mass moved must be g times as great as the quantity of matter used to define the unit force. Thus the unit of mass in the metric gravitational system is equal to 980.6 grams, and in the British gravitational system to 32.172 pounds.

Since the value of the attraction exerted by the earth on a given mass is not the same all over the surface of the earth, the gravitational unit

[1] The value of g is by no means the same at all points on the earth. (See § 116.)

of force will not everywhere be exactly equal to the force with which the earth attracts (*i.e.* the weight of) a gram or pound, as the case may be. For comparatively rough measurements, however, in which the change in the value of the gravitational attraction may be neglected, there is no doubt it is often convenient to use the gravitational units, particularly when, as is often the case, all the forces which have to be dealt with are due to the action of gravity on matter. For scientific purposes, however, there can be no doubt that the absolute system of units, in which mass is taken as the fundamental unit, is preferable.

The units of work or energy in the two gravitational systems are the work done in lifting a gram, or a pound, through a centimetre, or a foot, as the case may be, at the sea-level and at latitude 45°. These units are called the gram-centimetre, and foot-pound units of work respectively.

The following tables exhibit the connection between the fundamental units and the units of force, mass, and work in the absolute and gravitational systems :—

Metric System.

Fundamental Units.

Absolute.		Gravitational.	
Length	Centimetre	Length	Centimetre
Mass	Gram	Force	Weight of a gram at lat. 45° and sea-level
Time	Second	Time	Second

Derived Units.

Force	Dyne	Mass	980.6 grams
Work and Energy	Erg	Work and Energy	Gram-centimetre at lat. 45° and sea-level

British System.

Fundamental Units.

Absolute.		Gravitational.	
Length	Foot	Length	Foot
Mass	Pound	Force	Weight of a pound at lat. 45° and sea-level
Time	Second	Time	Second

Derived Units.

Force	Poundal	Mass	32.172 pounds
Work and Energy	Foot-poundal	Work and Energy	Foot-pound at lat. 45° and sea-level

Since the force with which the earth attracts a mass m is equal to mg absolute units of force, the work which is done when this mass is raised through a vertical height h is

$$mgh.$$

If m is expressed in grams, g in cm/sec²., and h in cm., *i.e.* if all the quantities are measured in *c.g.s.* absolute units, then the work mgh is expressed in ergs. If m is measured in pounds, g in foot/sec²., and h in feet, then the work mgh is expressed in foot-poundals.

If the experiment is performed at latitude 45° and at the sea-level, then the work done in raising m grams through a height of h centimetres is mh gram-centimetres. If, however, the experiment is performed at a point on the earth's surface where the acceleration of gravity is g, then the work done will be $mh.\frac{g}{980.6}$ gram-centimetres. For at the standard position the value of the acceleration of gravity is 980.6, and the force with which the earth attracts a given mass is proportional to the acceleration of gravity or g. In the same way the work done in raising m pounds through h feet at the standard position is mh foot-pounds, while at a place where the acceleration of gravity is g it is $mh\frac{g}{32.172}$ foot-pounds.

77. Graphic Representation of the Work done by a Force.— If a force (F) remains uniform while its point of application moves in the direction of the force through a distance s, the work done is $F.s$. Let the line OX (Fig. 56) represent space in the direction of the line of action of the force, and distances measured parallel to a line OY, at right angles to OX, represent the *magnitude* (not direction) of the force. If through M and N, any two points in OX, we draw two lines MP, NQ parallel to OY, and of such a length that they represent the magnitude of F, then the area of the rectangle MPQN will represent the work done by the force while its point of application moves through the distance MN.

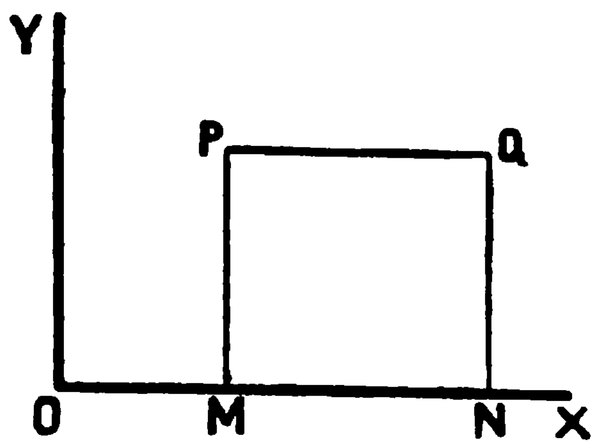


FIG. 56.

For the area of the rectangle is equal to $\overline{MN} \times \overline{MP}$, and MN represents s, and $\overline{MP}$ represents F, so that the product $F.s$ is represented by the area of the rectangle MPQN, that is, there are as many units of area in MPQN as there are units of work in $F.s$. Hence if a centimetre along OX represents a displacement of one centimetre, and a centimetre along OY represents a force of one dyne, then each square centimetre in the area will represent an erg.

If the force is not constant in magnitude during the displacement, we must, as was done in the case of speed in § 34, draw a curve such that the ordinate (MP) at any point M on the axis OX represents the force

acting when the point of application is at M. Then the area contained between any two ordinates, the curve and the axis OX, will represent the work done while the point of application of the force moves from the point corresponding to the foot of one ordinate to that corresponding to the foot of the other.

As an example of the application of this graphical method of representing the work done by a variable force, we may take the method employed to determine the work done by the steam during each stroke of a steam-engine piston. The curve is drawn mechanically by means of an instrument called an indicator; this consists of a small cylinder, the piston of which is held down by a spring. A pencil attached to this piston bears on a sheet of paper, which, by means of a connecting link, is moved backwards and forwards by the piston of the engine along a line at right angles to the direction in which the pencil is moved when the indicator piston is forced back. The cylinder of the indicator is connected by a pipe with the inside of the engine cylinder, so that the amount the spring is forced back, *i.e.* the movement of the pencil, is a measure of the pressure of the steam in the cylinder, and hence of the force acting on the engine piston. A curve obtained from such an indicator is shown in Fig. 57. Here the force acting on the piston is measured along OY, and the displacement of the piston along OX. The length MN represents the "stroke" of the piston, or the distance through which it travels. Starting from A, the part of the curve ABCD represents the pressure acting on the piston while it is moving in the direction of this pressure, and hence the work done *on* the piston is represented by the area MABCDNM. During the return stroke, however, which is represented by the part DEA of the curve, the piston is moving against the pressure, and the work done *by* the piston is represented by the area MAEDNM. Hence the work done *on* the piston during a complete to-and-fro motion is the difference between these areas, namely, the area ABCDEA.

TRAVEL OF PISTON IN INCHES

Fig. 57.

By counting up, it will be found that the figure ABCDEA contains about 230 small squares. Now one side of each of these squares represents a pressure of 4 lbs. per square inch on the piston, while the other side represents a travel of the piston through a space of one inch. Each square will therefore correspond to 4/12 or 1/3 foot-pound of work.

Hence the work done by the steam on each square inch of the piston during a whole revolution of the crank is 230/3, or 76.8 foot-pounds. If s is the area of cross section of the piston, then the work done during one revolution is 76.8 s foot-pounds.

78. Power or Activity.—It will be noticed that the work done by a force during a given displacement is independent of the time taken. Thus the same amount of work is done against gravity by an engine which drags a train up a given incline in an hour, as would be done if it had done this in a minute. Since there is practically a very great difference between these two cases, it is obvious that we have to take the time an agent takes to do a certain quantity of work into account. The rate of doing work, or the work done divided by the time taken to do it, is called the *Power* or *Activity* of an agent. Hence in the above case the power would in one case be sixty times greater than in the other.

An agent which is capable of raising 550 pounds through a vertical distance of one foot[1] (*i.e.* of doing 550 foot-pounds) in one second is said to be of one *horse-power*. The French horse-power (*cheval-vapeur*) is such that 75 kilograms can be raised one metre in each second.

The practical unit of power based on the *c.g.s.* system is one joule per second, and is called the watt. Hence a watt is 10^7 ergs per second.

One horse-power is equal to 746 watts.

79. Energy.—We find by experience that under certain circumstances bodies are capable of doing work. Thus when a weight has been raised up above the surface of the earth, it possesses the capacity for doing work during its return to the surface of the earth. For a string attached to the weight may pass over a pulley, and the other end be attached to a bucket of water. If the bucket of water is lighter than the weight, then during its fall it will raise up the bucket, and hence do work. Again, a body which is in motion possesses the power of doing work while it is losing its motion, as, for instance, a bullet when it strikes a block of wood loses its motion, but in doing so it does work on the block, for it penetrates the block against the resisting force of the wood. Hence we see that under certain circumstances bodies possess a capacity for doing work; and this capacity for doing work is called *Energy.*

The energy of a body is measured by the work it can do while changing to some standard state, or, what is sometimes more convenient, the work which has to be done on the body to bring it from some standard state to the actual state. Thus in the above examples we may measure the work the weight can do before it comes to rest on the surface of the ground, or the work the bullet can do before it is brought to rest.

[1] To be strictly accurate, this is only true at latitude 45° and sea-level; however, for most engineering purposes, the small differences in the value of g at different places do not cause any appreciable errors.

It will be noticed that in the two examples given the nature of the circumstance owing to which the body possesses energy are different. In the first case, that of the raised weight, it possesses energy due to its position relative to the surface of the earth; while in the case of the moving bullet, the energy possessed is due to the motion. The energy of a body which is due to its motion is called *Kinetic Energy,* while the energy due to its position and not to its motion is called *Potential Energy.* Thus the raised weight possesses potential energy and the moving bullet kinetic energy.

80. Potential Energy.—Examples of potential energy abound in everyday life; thus when a clock-weight is raised we do work against the attraction which exists between the weight and the earth. The raised weight, however, possesses potential energy, and by the loss of this it keeps the clock in motion, doing work in overcoming the friction of the parts of the clock, imparting motion to the air, and thus enabling us to hear the sound of the tick, &c. When winding up a watch, in the same way, work is done in bending the spring, and its energy is increased so that it possesses potential energy, and can do work while it unbends.

In the case of the bent spring, or in that of a piece of stretched india-rubber cord, it is evident that the material of the spring or cord is in a state of strain, and it is owing to this strain that the body possesses potential energy. In the case of the raised weight, however, we are unable to detect anything connecting the weight and the earth which is strained and which tends to revert to its former state, and thus endows the raised weight with its potential energy. If we were able to state what was the nature of the stress in this case, we should know the *cause* of gravitation. We are, however, unable to do this, and must content ourselves with saying that one portion of matter exerts force on another, and that it is to this force that their mutual potential energy when separated is due (compare § 64). We shall find that the theory that there is an all-pervading massless medium is necessary to explain the observed facts in heat, light, and magnetism and electricity; we are not at present, however, entitled to say that it is to stresses set up in *this* medium that gravitation is due.

81. Kinetic Energy.—Suppose that a body of mass m is moving with a speed v, and that a force F acting on the body in an opposite direction to that of the motion brings the body to rest after it has passed over a distance s. Then the work done by the body on the force while coming to rest is Fs. Now if in a distance s the body loses a speed v, the acceleration must be $a=v^2/2s$ (§ 35). Also the force F produces this acceleration in the mass m, and hence $F=ma=mv^2/2s$. But the work done by the force is Fs, that is, $\frac{1}{2}mv^2$. Hence the kinetic energy of a body of mass m moving with a speed v is $\frac{1}{2}mv^2$.

82. Change of Form of Energy.—The energy of a body is capable

of changing its form from potential to kinetic, and *vice versa*. Thus suppose a stone of mass m is supported on the edge of a cliff at a height h above the base of the cliff. The potential energy is equal to the work done in raising the stone through a vertical height h. The force with which the earth attracts the stone is equal to the product of the mass of the stone (m) into the acceleration which the force would produce (*i.e.* g). Hence in raising up the stone it has been moved through a space h against a force mg, and therefore the work done has been mgh, so that this is its potential energy.

If now the stone be allowed to fall freely, it will gradually lose its potential energy, but will at the same time acquire velocity and hence kinetic energy. After it has fallen a distance s, its speed will be given by the equation $v^2 = 2gs$, and hence its kinetic energy ($\frac{1}{2}mv^2$) will be equal to mgs. The potential energy is now $mg(h-s)$, since the stone is now at a height $(h-s)$ above the ground. The sum of the kinetic and potential energies is therefore equal to $mgs + mg(h-s)$ which is equal to mgh, the original potential energy. When the stone reaches the ground its potential energy is zero, but the speed which it has acquired is now given by $v^2 = 2gh$, and hence the kinetic energy ($\frac{1}{2}mv^2$) is equal to mgh, the same value as the potential energy at the start.

Thus, although during the fall of the body there is a gradual change of potential energy into kinetic, the total energy remains constant.

If the stone were thrown vertically upwards with a speed v, then to start with the kinetic energy would be $\frac{1}{2}mv^2$. This would gradually diminish as the stone rose and lost speed; there would, however, be a gain of potential energy. When the stone is at the top of its flight it comes for an instant to rest, so that its kinetic energy is now zero. It however possesses potential energy exactly equal in amount to its original kinetic energy.

83. Principle of the Conservation of Energy.—In the previous section we saw that although the *form* of the energy of the stone altered, so that it existed sometimes as potential energy, sometimes as kinetic energy, and sometimes as both together, yet the total quantity of energy was constant throughout. We shall in subsequent sections consider many other forms in which energy can exist besides those already considered, and we shall find that these different forms of energy can be converted from one kind to the other.

When the stone reaches the ground it is brought to rest by impact with the ground, when it will apparently have lost both its kinetic energy and its potential energy. It is, however, found that the energy has not been *lost*, but has been transformed into another form, namely, that of heat, so that both the stone and the part of the earth it struck are warmer than they were before the impact. Joule has indeed shown, as we shall see later (§ 251), that in every case a given amount of work entirely spent in producing heat always produces the same quantity of heat, no matter

how the conversion takes place. For instance, a given number of ergs of work which can be obtained from stopping a moving bullet, and hence destroying its *kinetic* energy, will produce a certain quantity of heat. The same number of ergs of work done in rubbing a button on a piece of wood will produce exactly the same quantity of heat. The energy in the form of heat possessed by a body is supposed to be due to the motion of its particles. If this be so, then the kinetic energy of the stone moving as a whole is converted, by impact with the ground, into kinetic energy of the *particles* of the body and the earth near the point of impact, the particles moving (probably backwards and forwards with a vibratory motion) with, on the whole, a greater velocity than before.

These observations may be generalised, for in every case, without exception, it is found that the sum total of all the energy within any given boundary, through which energy is not allowed to pass, remains constant, although the energy within the boundary may be transformed into any of the many forms in which it is capable of existing.

The above statement amounts to an enunciation of a doctrine which is practically the keystone of modern science, and is known as the *doctrine or principle of the conservation of energy.*

It follows that if the boundary considered includes the universe, the principle of the conservation of energy amounts to a statement that the sum total of the energy of the universe is a fixed unalterable quantity.

The principle of the conservation of energy also denies the possibility of "perpetual motion." By "perpetual motion" is meant the devising of some arrangement so that energy in the form of mechanical work could be produced without energy in some other form being used up by the machine. Thus if an engine could be made to do work on external bodies for an indefinite time, and thus give out energy, without being supplied with energy from without, or diminishing the stock of energy in all its various forms which it originally possessed, we should have a means of creating energy, and this is in direct contradiction to the principle of the conservation of energy.

84. Availability of Energy.—Although the total quantity of energy in the universe remains a constant quantity, so that whenever a given amount of energy in any one form disappears an exactly equal quantity of energy in some other form makes its appearance, yet there is, as far as man is concerned, a very important difference between the states of energy. Thus, to recur to the illustration of the stone on the edge of a cliff. By a suitable mechanical arrangement we can utilise some portion at any rate of the potential energy of the stone, and make the stone do work. For example, by attaching a string to the stone we can make it keep a clock going, and strike the hours on a bell. Again, we can let the stone fall, and utilise the kinetic energy it possesses by making it do work, say in driving in a pile. Suppose, however, it is simply allowed to fall on the ground. As a result of the impact with the ground its kinetic energy

is destroyed, and the stone and the earth become warmer, and it is this heat which represents the lost kinetic energy; but in a very short time the heat of the stone, &c., will have diffused itself amongst surrounding objects. Energy in this form, that is, of uniformly diffused heat, is unavailable to man for the purpose of doing work; it is only when we have a body which is hotter than surrounding objects that we can utilise its heat energy to do work. Thus as far as we are concerned the energy of the stone has been wasted, for it is no longer available for doing work.

We shall find that in every transformation of energy there is some of the energy converted into heat, which becomes diffused throughout the universe, and is wasted. Thus, in the case of the raised stone being utilised to drive a clock, the friction at all the bearings causes some of the energy to be converted into heat. Heat also is produced each time the escape-wheel strikes the pallet of the detent, as well as the sound which we call the tick, and which we shall see is simply due to motion of the particles of the air, and this motion is gradually frittered away into heat by friction between the particles and other causes. Hence in this case all the energy of the raised stone eventually becomes changed into diffused heat, although in the meantime it goes through many transformations. Since in every transformation of energy from one form to another some of the energy becomes converted into uniformly diffused heat, the total quantity of *available* energy of the universe is continually diminishing.

This continual degradation of energy, which accompanies every phenomenon with which we are acquainted, leads us to two conclusions: Firstly, since the quantity of unavailable energy is continually increasing, there must have been a time when none of the energy of the universe was unavailable, and before which no phenomenon, such as we are acquainted with, can have occurred, for every such phenomenon necessarily involves a degradation of energy. Secondly, there must necessarily arrive a time when all the energy will be unavailable, the whole universe having become a uniformly hot, inert mass.

85. Energy of Rotation.—If a particle of mass m is moving in a circle of radius r, with angular velocity ω, then the speed of the particle is ωr; hence the kinetic energy is $\frac{1}{2}m\omega^2r^2$. We may suppose that a rotating rigid body is made up of a number of particles of mass m_1, m_2, m_3, &c., at distances r_1, r_2, r_3, &c., from the axis of rotation. If ω is the angular velocity of the body, then the speeds of the particles are ωr_1, ωr_2, ωr_3, &c., since ω is the same for all. Hence the kinetic energy (E) of the rotating body, being the sum of the energies of all the particles, is given by the equation—

$$E=\frac{1}{2}m_1\omega^2r_1^2+\frac{1}{2}m_2\omega^2r_2^2+\frac{1}{2}m_3\omega^2r_3^2+, \text{ \&c.},$$

or, since ω is common to all—

$$E=\frac{1}{2}\omega^2\{m_1r_1^2+m_2r_2^2+m_3r_3^2+, \text{ \&c.}\}$$

The quantity within the brackets, which is the sum of the product of the mass of each particle of the body into the square of its distance from the axis of rotation, is called the *Moment of Inertia* of the body. Hence, if K is the moment of inertia, we have

$$E=\tfrac{1}{2}K\omega^2.$$

The value of the moment of inertia for some of the simpler forms is given in the following table, in which M is in each case the mass of the rotating body :—

MOMENTS OF INERTIA.

Figure.	Position of Axis about which Rotation takes place.	Moment of Inertia.	
Solid cylinder .	Perpendicular to axis of cylinder, and passing through its mid point	$M\left(\frac{l^2}{12}+\frac{r^2}{4}\right)$	l=length r=radius
Solid cylinder .	Coincident with axis of cylinder	$M.\frac{r^2}{2}$	r=radius
Hollow cylinder	Coincident with axis of cylinder	$M\left(\frac{R^2+r^2}{2}\right)$	R=external radius r=internal radius
Sphere . . .	Passing through centre	$M.\frac{2r^2}{5}$	r=radius
Rectangular bar	Through centre of two opposite faces	$M\left(\frac{a^2+b^2}{12}\right)$	a and b are adjacent edges of the faces through which the axis passes

It will be noticed that in every case the moment of inertia is equal to the product of the total mass of the body into an expression involving the geometrical form of the body. This suggests another way of obtaining an expression for the energy of a rotating body. If the whole mass M of the body were concentrated at a single point which moved in a circle of radius R with the same angular velocity as the body, then the energy due to the rotation would be $\frac{1}{2}M\omega^2R^2$. If R is so chosen that the energy is the same as is actually the case, then

$$\tfrac{1}{2}K\omega^2=\tfrac{1}{2}M\omega^2R^2.$$

$$\therefore \quad K=MR^2.$$

The value of R, which satisfies this equation, is called the *radius of gyration* of the body. Hence the radius of gyration may be defined as the distance from the axis of rotation, at which the whole mass of the

body must be supposed concentrated, in order that the energy of rotation may be the same as it is actually. The value of the square of the radius of gyration for each case given in the table of moments of inertia is obtained by omitting the factor M in the expressions for the moment of inertia.

86*. Impact of Inelastic Bodies.—Suppose a body of mass m_1, moving with a speed v_1, strike another body of mass m_2 which is at rest, and the two bodies adhere together after impact. If v is the speed with which the two bodies move after impact, then their momentum is $(m_1+m_2)v$. Before impact the momentum of the first body is $m_1 v_1$, and that of the second zero. Hence, since the total momentum of the two bodies must be the same before and after impact, we have

$$(m_1+m_2)\,v=m_1v_1.$$

$$\therefore \qquad v=\frac{m_1}{m_1+m_2}\,v_1.$$

If the second body is itself moving with speed v_2 in the same straight line as the first, the momentum before impact is $m_1v_1+m_2v_2$. Hence

$$(m_1+m_2)\,v=m_1v_1+m_2v_2.$$

$$\therefore \qquad v=\frac{m_1v_1+m_2v_2}{m_1+m_2}.$$

If the bodies are originally moving in the same direction, then v_1 and v_2 are of the same sign, and therefore the motion after impact is in the same direction as before impact. If, however, the bodies are originally moving in opposite directions, so that v_1 and v_2 are of opposite sign, v will be of the same sign as the greater of the two products m_1v_1, m_2v_2. Hence the direction of motion after impact will be that of the motion of the body having the greater *momentum.*

In the case where one of the bodies is originally at rest, the kinetic energy E before impact is given by

$$E_1=\tfrac{1}{2}\,m_1v_1^2,$$

and the kinetic energy E_2 after impact by

$$E_2=\tfrac{1}{2}\,(m_1+m_2)\,v^2.$$

Substituting in this last expression the value of v, we have

$$E_2=\tfrac{1}{2}\,(m_1+m_2)\left(\frac{m_1}{m_1+m_2}\right)^2 v_1^2$$

$$=\tfrac{1}{2}\frac{m_1^2v_1^2}{m_1+m_2}=\tfrac{1}{2}m_1^2v_1^2\div(m_1+m_2).$$

Hence $$E_2=\frac{E_1}{m_1+m_2}.$$

The kinetic energy of the two bodies is therefore less after impact than before. The reason for this is that some of the kinetic energy has been

transformed into heat energy by the impact. In the same way, in the case when the two bodies are originally in motion, the sum of the kinetic energies of the bodies before impact is greater than the kinetic energy of the combined bodies after impact, some of the kinetic energy having been converted into heat energy.

87*. Impact of Elastic Bodies.—In the case of elastic bodies, after impact they will separate, in general moving with different velocities, the total momentum before impact being, as in the case where the bodies are inelastic, equal to the total momentum after impact.

Newton proved by experiment that if two spheres of the same material collide, so that their centres are moving along the same straight line, then their relative speed before impacts bears a constant ratio to their relative speed after impact. If u_1 and u_2 are the velocities before impact, and v_1 and v_2 the velocities after impact, then the relative velocity of the spheres before impact is $u_1 - u_2$, and that after impact $v_1 - v_2$. Newton's results show that the ratio $\frac{v_1 - v_2}{u_1 - u_2}$ is a quantity independent of the mass of the spheres, at any rate within fairly wide limits, or the velocity with which they are moving; it simply depends on the material of which they are composed. This ratio is always a negative quantity, never greater, and in most cases less, than unity. This means that the relative velocity after impart is in most cases *less*, and cannot be greater, than the relative velocity before impact; also that if u_1 is greater than u_2, then v_1 is less than v_2, so that $v_1 - v_2$ is of opposite sign to $u_1 - u_2$.

Hence v_1 being greater than v_2 means that the body moving with velocity v_1 "strikes" the other, but that this latter after the impact will now move with the greater velocity.

The constant fraction $\frac{v_1 - v_2}{u_1 - u_2}$, which depends simply on the material of which the spheres are composed, is called the *Coefficient of Restitution* of the material. If the coefficient of restitution is unity, then the relative velocity of the bodies is unchanged by the impact, and the material of which they are composed is said to be perfectly elastic.

If the coefficient of restitution is e, then

$$-\frac{v_1 - v_2}{u_1 - u_2} = e.$$

Also, since the momentum before and after impact is the same,

$$m_1 u_1 + m_2 u_2 = m_1 v_1 + m_2 v_2.$$

Solving these two equations for v_1 and v_2, we get

$$v_1 = \frac{u_1(m_1 - e m_2) + m_2 u_2(1 + e)}{m_1 + m_2},$$

and

$$v_2 = \frac{u_1 m_1(1 + e) + u_2(m_2 - e m_1)}{m_1 + m_2}.$$

It will be noticed that if $e=o$, *i.e.* if the balls are inelastic and adhere together after impact, we have

$$v_1=v_2=\frac{u_1m_1+u_2m_2}{m_1+m_2},$$

the same value as that obtained in the previous section.

In the case of the direct impact of a sphere on a fixed surface, we may obtain the velocity of rebound by making u_2 zero and m_2 infinite in the above expressions. Dividing numerator and denominator by m_2, we get

$$v_1=\frac{u_1\left(\frac{m_1}{m_2}-e\right)+u_2(1+e)}{\frac{m_1}{m_2}+1}.$$

When m_2 is infinite, m_1/m_2 is zero. Hence, if u_2 is zero,

$$v_1=-u_1e.$$

Therefore the velocity of rebound is in the opposite direction, and equal to e times the velocity before impact. Since e is less than unity, the speed after impact is less than before. This result might have been arrived at directly from Newton's experimental result, for the relative velocity of the sphere and the plane before impact was u_1, and after impact it is v_1; hence

$$-\frac{v_1}{u_1}=e, \text{ or } v_1=-eu_1.$$

It will be found, on calculating the value of the kinetic energy before and after impact, that, except in the case where $e=1$, the total kinetic energy after impact is less than before, some having been transformed into heat. When $e=1$, the total kinetic energy before and after impact is the same.

88*. Oblique Impact.—Hitherto we have only considered the case where the two spheres, before impact, were travelling, either in the same or in opposite directions, along the same straight line. When the impact is oblique, we resolve the velocity of each sphere along and perpendicular to the line joining the centres of the spheres at the instant of contact. If the spheres be smooth, then there will be no force exerted at right angles to the line joining the centres, and hence the components of the velocities along this line will be unaltered by the impact. The components along the line joining the centres will obey the laws given for direct impact in the last section, and hence the components in this direction after impact can be obtained. By compounding these components with those which remained unaltered, the velocity of the spheres after impact can be obtained.

In the case of a sphere, such as a billiard-ball, striking the cushion AB (Fig. 58) with a speed u, and at an angle θ to the normal to the cushion, the component of the velocity parallel to the cushion is $u \sin \theta$, and that perpendicular to the cushion is $u \cos \theta$. The first of these components remains unaltered, and the other becomes $-eu \cos \theta$ (§ 87), where e is the coefficient of restitution. Hence we have to compound a velocity of $u \sin \theta$ parallel to $\overrightarrow{BA}$ (Fig. 58) with a velocity of $eu \cos \theta$ parallel to $\overrightarrow{CN}$. If the resultant velocity is v, and makes an angle ϕ with the normal CN, we have, since the components are at right angles—

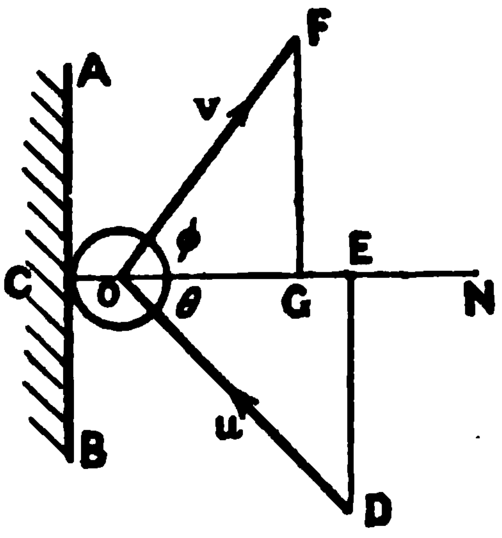


FIG. 58.

$$v^2 = OF^2 = \overline{GF}^2 + \overline{OG}^2 = u^2 \sin^2 \theta + e^2 u^2 \cos^2 \theta.$$

$$\text{Also} \quad \frac{u \sin \theta}{eu \cos \theta} = \frac{\overline{GF}}{\overline{OG}} = \tan \phi.$$

$$\therefore \qquad e \tan \phi = \tan \theta.$$

Hence the ball rebounds so as to travel more nearly parallel to the cushion than before impact, for e is less than unity.

If $e = 1$, then $v = u$, $\tan \phi = \tan \theta$, and therefore $\phi = \theta$, so that the ball wound rebound with the same speed, and with its direction of motion inclined to the cushion at the same angle as before, but on the opposite side of the normal.

CHAPTER XI

MACHINES

89. Simple Machines.—A contrivance by means of which a force applied at one point gives rise to a force, often different in magnitude and direction to the impressed force, at some other point is called a machine. A steam-engine, a hydraulic press, or a testing-machine, are all examples of machines; they, however, consist of a number of separate parts, each of which satisfies the definition of a machine given above, and which may be classified under various heads. Each of these separate parts is called a *simple machine*.

In studying these simple machines, we shall suppose that the machine is in equilibrium, so that the force impressed at one point is just balanced by the force impressed at some other point. One of these forces, which is impressed on the machine by some other body, is generally called the *Power*, while the other is called the *Weight*. It must be carefully borne in mind that the term "power" in this relation has no connection with the same word used in § 78 to denote the rate of doing work, nor does the term "weight" necessarily mean that the machine is used to raise or support a mass against the attraction of gravity.

In finding the conditions which have to be fulfilled, in order that the simple machines may be in equilibrium, we suppose that they work without friction. Under these circumstances, if we suppose the machine to receive a small displacement of such a nature that the connection between its parts is not in any way disturbed, the work done by the "power" must, according to the principle of the conservation of energy, be equal to the work done on the "weight." Hence, if P and W are the two forces which balance one another on a machine, and when the machine receives a small displacement, if the displacement of the point of application of P parallel to its line of action be p, and that of W parallel to its line of action be w, then the work done by P is Pp, and the work done on W is Ww. Hence

$$Pp = Ww,$$

or

$$\frac{P}{W} = \frac{w}{p}.$$

From this we see that the displacements are *inversely* as the forces, so that if a small "power" is to exert a large "weight," the displacement of the "power" must be large compared to the displacement of the

"weight." This is often expressed by saying that "what is gained in power is lost in speed."

90. The Lever.—A lever is a rigid bar, either straight or curved, which is capable of a motion of rotation about a fixed point, called the fulcrum.

Since the lever, when in equilibrium, is under the action of three forces—the "power," the "weight," and the reaction of the fulcrum, it follows (§ 73) that the lines of action of all these forces must lie in one plane, and either be parallel or meet at a point.

The most direct way of obtaining the relation between the "power" and "weight" in the lever is to take moments round the fulcrum. If the lever is to be in equilibrium, these moments must be equal and opposite. Hence if a is the perpendicular distance between the fulcrum and the line of action of the "power," and b that between the fulcrum and the line of action and the "weight," we have

$$Pa = Wb,$$

or

$$\frac{P}{W} = \frac{b}{a}.$$

In the case when the lines of action of the forces are at right angles to the lines joining the points of application to the fulcrum, a and b represent the distances of the points of applications of the forces from the fulcrum, and are called the *arms* of the lever.

It is usual to divide levers into three classes, according to the relative positions of the points of application of the forces and the fulcrum. In Class I., (*a*), Fig. 59, the fulcrum F lies between the points of application, A and B, of the forces. In this case the "power" may, as in the crowbar, be applied at the end of a longer arm, and so be used to exert an

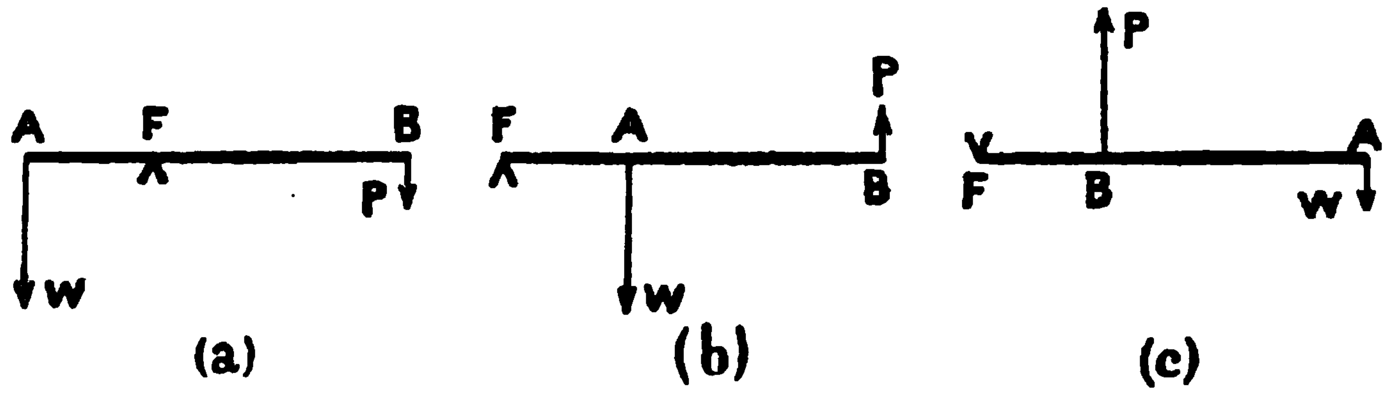


Fig. 59.

increased force; or it may, as in the catapults used by the ancients for throwing heavy stones, act at the end of the shorter arm, so that the distance through which the point of application of the "weight" moves is greater than that through which that of the "power" moves. Other examples of levers of this class are the beam of a balance, a pair of scissors, and a pump handle.

In a lever of the second class, (*b*), Fig. 59, the point of application A

of the "weight" lies between the point of application B of the "power" and the fulcrum F. In this class the "power" is always less than the "weight." Examples of levers of the second class are nut-crackers, and an oar. In this latter case the fulcrum is the blade of the oar, which remains approximately at rest in the water, and the pressure on the row-lock is the "weight."

In the third class of levers, (*c*), Fig. 59, the "power" is applied at a point between the fulcrum and the point of application of the "weight," so that the "power" must always be greater than the "weight." An example of a lever of this class is an ordinary pair of tongs. The fore-arm is another example, the fulcrum being at the elbow-joint, and the power applied a few inches along the fore-arm, where the biceps muscle is inserted in the radius.

In every case the reaction of the fulcrum will be equal to the resultant of the "power" and the "weight," and may be obtained by the methods already given for compounding two forces, whether meeting at a point or parallel, as the case may be (§§ 66 and 69).

91. The Wheel and Axle.—This machine consists of two drums or wheels of different diameters fixed to the same axle. A rope coiled round the drum of smaller diameter carries the "weight," while another rope coiled round the other drum, but in an opposite direction, carries the "power." It will be seen from Fig. 60 that the arrangement is virtually a lever with the fulcrum at the axis B, about which the drums can turn. Hence $P \times \overline{BC} = W \times \overline{AB}$.

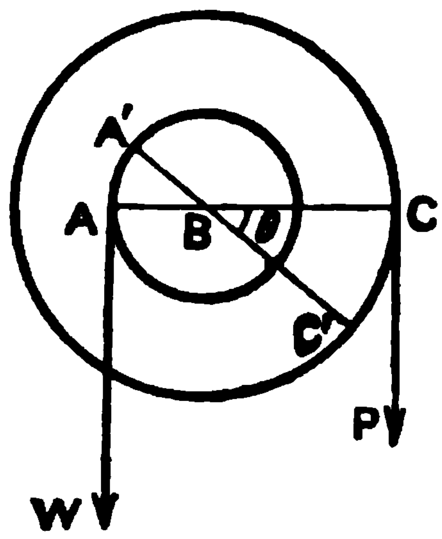


FIG. 60.

This relation may also easily be found from the principle of work; for suppose the two drums to turn through an angle θ, so that the point C comes to C′ and A to A′. Then the length of rope coiled up on the smaller drum (the axle) is equal to the arc AA′, and this represents the distance through which the point of application of W has been moved. Hence the work done on W is $W \times \overline{AA'}$. But the arc AA′ is equal to $\overline{AB}.\theta$, where θ is the circular measure of the angle through which the machine has turned. Therefore the work done on W is $W.\overline{AB}.\theta$. In the same way the work done by P is $P.\overline{BC}.\theta$.

Hence

$$W.\overline{AB}.\theta = P.\overline{BC}.\theta$$

$$\therefore \quad W.\overline{AB} = P.\overline{BC}$$

or

$$\frac{W}{P} = \frac{\overline{BC}}{\overline{AB}}.$$

The principle of the wheel and axle is used in the capstan and in the windlass. In these arrangements the "power," instead of being applied

to the wheel by means of a rope wound round the circumference, is applied to the end of one or more rods which virtually form spokes of the wheel, the direction of the line of action of the "power" being continually altered, as the machine turns round, so as to be always at right angles to the spoke. If the "weight" and "power" bear to one another exactly the ratio given above, the machine will be in equilibrium; if P exceeds this value there will be rotation, and W will be drawn up.

92. The Pulley.—A pulley consists of a disc or wheel, called the sheaf, mounted on an axle which is fixed to a framework called a block. The edge of the disc is usually grooved so that a cord can lie round it. If the block is fixed, then the direction of a force, but not its magnitude, may be changed by means of a pulley. If a tension P be applied to one end of a string which passes over such a pulley, then, since if we neglect the friction of the pulley and the stiffness of the string the tension is the same throughout the string, in order to keep the pulley in equilibrium the other end of the string must be pulled with a force P. This is at once evident, for if A and C (Fig. 61) are the points where the string leaves the pulley, and B is the centre of the sheaf, then taking moments about B,

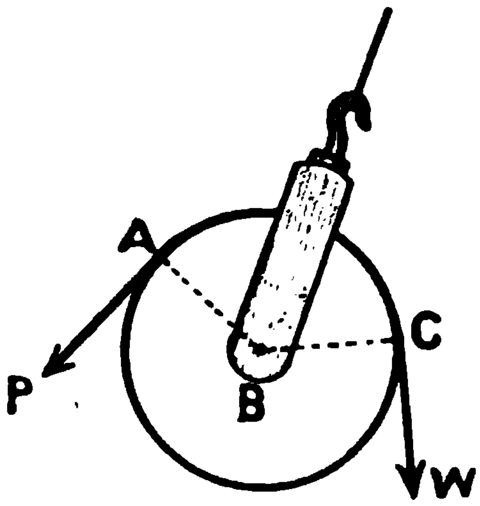


FIG. 61.

$$P \times \overline{AB} = W \times \overline{BC}.$$

$$\text{But } \overline{AB} = \overline{BC}. \quad \text{Hence } P = W.$$

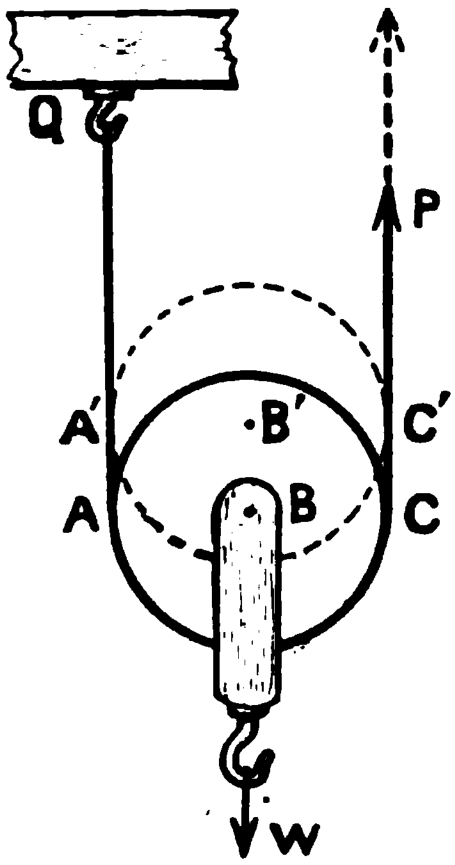


FIG. 62.

If the block of the pulley, instead of being fixed, is attached to the weight, while one end (Fig. 62) of the string is attached to a fixed support, while the "power" acts at the other end, P may be less than W. If, as is usually the case, the two portions of the string QA and CP are parallel, and the pulley moves through a distance h from the position ABC to the position A′B′C′, then the end of the string where P is attached will move up through a distance $2h$, for the portion QA of the string has been shortened by a length h, and the point C has also risen through a height h. Therefore, while W has been raised through a height h, P must have moved through a distance $2h$, so that, equating the work done in the two cases,

$$hW = 2hP$$

or

$$W = 2P.$$

In this expression, since the pulley itself has to be raised, we must include its weight in W. Otherwise, if W is the weight of the pulley and W' is the weight supported, we have

$$W' + W = 2P.$$

There are several arrangements in which more than one pulley is used, but we shall only describe one of these, which is the only one that is used in practice. It consists of two blocks, each fitted with several sheafs, which usually all turn on the same axle. One of the blocks is attached to a fixed point, while the other is attached to the "weight." One end of the string is attached to one of the blocks, it then passes round one of the sheafs of the other block, then over one in the first block, and so on till it has passed over all the sheafs. If the string passes n times from one block to the other, then we shall have

$$W = nP,$$

where W includes the weight of the movable block.

It will be seen in Fig. 63, where there are three sheaves in the movable block and the string passes six times from one block to the other, that if the movable block, and therefore also W, is raised through a height h, then the free end of the string will have to move through a distance $6h$. Hence the work done on W is hW, and that done by P is $6hP$. Therefore

FIG. 63.

$$hW = 6hP$$

or

$$W = 6P.$$

93. The Inclined Plane.—Suppose a body G (Figs. 64 and 65) rests on an inclined plane AB, and that there is no friction between the body and the surface of the plane, or at any rate that by suitable means friction is so much reduced as to be negligable in comparison with the other forces in play. The weight of the body W acts vertically downwards, and the reaction of the plane acts perpendicular to the surface AB, so that if the body is to be prevented from sliding down the plane under the influence of the resultant of these two forces, it must be acted upon by a third force. The two principal cases which occur are when this force acts along a

direction parallel either to the hypotenuse AB (Fig. 64) or the base AC (Fig. 65). In both cases the most convenient method of obtaining the relation between the force P, which we may call the "power," and W, for any given plane, is to use the principle of work.

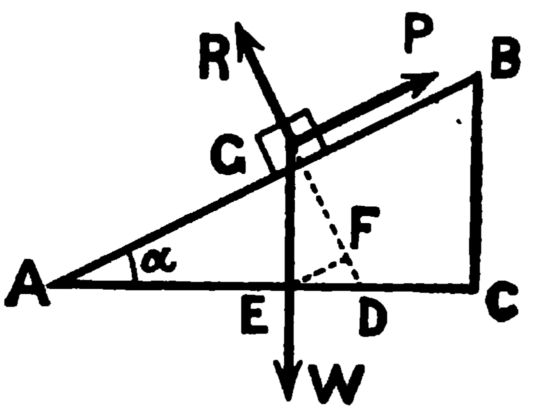


FIG. 64.

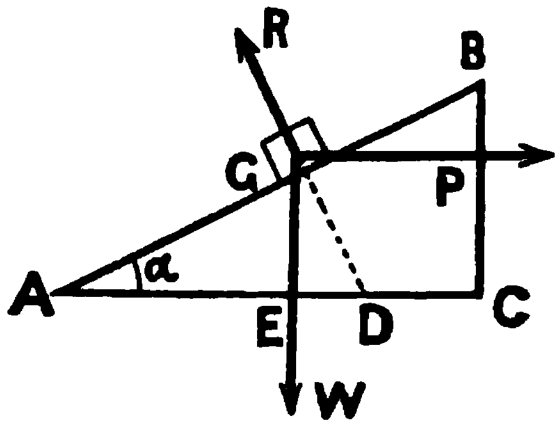


FIG. 65.

Suppose G to move, under the influence of P, from A to B. Then the work done by P is $P.\overline{AB}$, if P acts parallel to the hypotenuse, and $P.\overline{AC}$ if P acts parallel to the base, for in the one case the point of application has moved through a distance $\overline{AB}$ in the line of action of the force, while in the other case the component of the displacement in the line of action is $\overline{AC}$. In both cases the work done on W is $W \times \overline{CB}$, since the component of the movement of the point of application of W, parallel to its line of action, is $\overline{CB}$. As in both cases the point of application of R moves at right angles to its line of action, no work is done on or by R during the displacement. Hence we have, when P acts parallel to $\overline{AB}$,

$$P.\overline{AB} = W \times \overline{CB}$$

$$\frac{P}{W} = \frac{\overline{CB}}{\overline{AB}}.$$

When P acts parallel to $\overline{AC}$,

$$P.\overline{AC} = W \times \overline{CB}$$

$$\therefore \quad \frac{P}{W} = \frac{\overline{CB}}{\overline{AC}}.$$

The above relations may also be easily found by resolving W along the direction of the line of action of P, and perpendicular to the surface of the plane. The first of these components will be equal and opposite to P, and the other will be equal and opposite to R. The triangle of forces may also be employed, for if the direction of R be produced back to D and we draw EF perpendicular to GD, then the triangle GEF in Fig. 64, or the triangle GED in Fig. 65, has its sides parallel to the three forces; and hence these forces are proportional to the sides of the triangle, and this triangle is, in each case, similar to the triangle ABC.

The further working out of these two cases will form a useful exercise for the reader.

94. The Screw.—If a right-angled triangle ABC (Fig. 66) cut out of paper be wrapped round a cylinder, so that the base AB of the triangle lies entirely in a plane at right angles to the axis of the cylinder, then the hypotenuse AC will trace out a spiral line *aidfgc* on the surface of the cylinder. If $\overline{AM}$ is equal to the circumference of the cylinder, then the distance, measured parallel to the axis of the cylinder, between the point *a*, which corresponds to the apex A of the triangle, and the point *e*, where

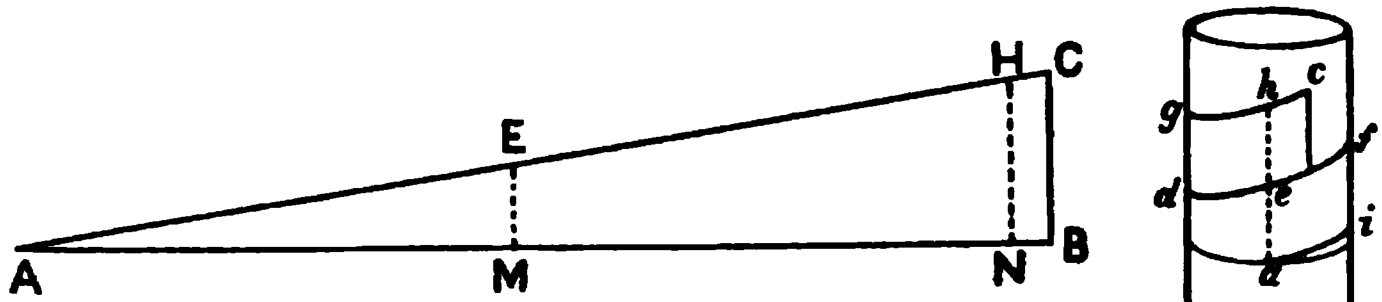


FIG. 66

a line drawn through *a* parallel to the axis meets the spiral line, is equal to $\overline{EM}$. Hence if starting at *a* we follow the spiral line for a complete turn, we shall move along the cylinder parallel to the axis through a distance $\overline{ae}$ equal to $\overline{ME}$. If we go twice round we shall move through a distance $\overline{ah}$ equal to $\overline{NH}$, where $\overline{MN}$ is equal to $\overline{AM}$. But since $\overline{AN}$ is double $\overline{AM}$, $\overline{NH}$ must be double $\overline{ME}$. Hence $\overline{ah}$ is double $\overline{ae}$, so that for every complete turn the spiral line advances parallel to the axis through an equal distance. If a projecting ridge were fixed to the outside of the cylinder along the spiral line we should have a *screw*, the projection forming the *thread*. The distance between two consecutive threads is called the *pitch* of the screw. If a hollow cylinder has a groove cut on its inside surface so as just to fit the screw, it forms a *nut*. If the nut is turned through 360°, or one whole turn, it will move along the screw through a distance equal to the interval between two consecutive threads, or to the pitch of the screw.

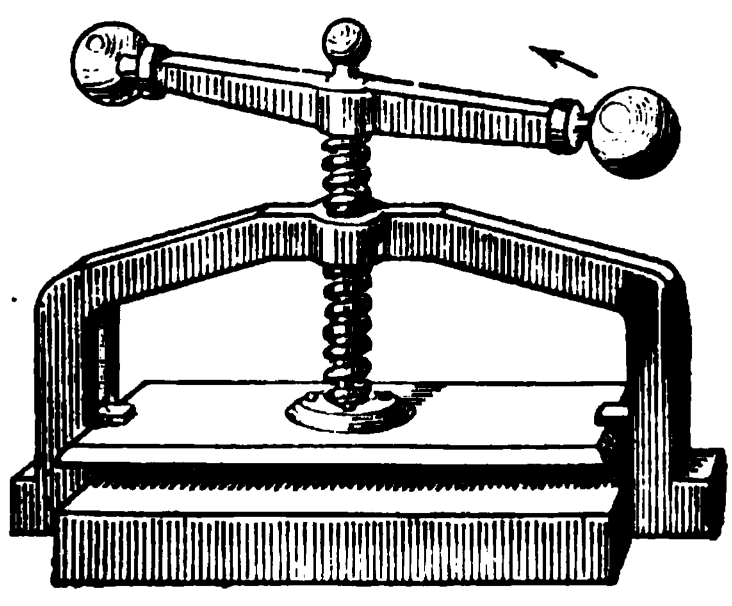


FIG. 67.

If, as in a screw press, we have a force P acting at right angles to the end of a cross arm (Fig. 67) of length $2r$ attached to a screw of which the pitch is h, so that the distance of the point of application of P from the axis of the screw is r, then during a complete turn the point of application of P will move through a distance equal to the circumference of a

circle of radius r, that is, $2\pi r$; and hence, since P is supposed always to act at right angles to the cross arm, the work done by P is $2\pi rP$. If W is the force exerted by the screw parallel to its axis, then the work done on W during a complete turn is hW, since the point of application of W will have been driven back through a distance equal to the pitch of the screw. Hence

$$2\pi rP = hW$$

or

$$\frac{P}{W} = \frac{h}{2\pi r}.$$

For a given value of P we see that W is increased by decreasing the pitch of the screw, *i.e.* by having more threads to the inch, and by increasing the leverage r at which P acts.

In deducing the above formula we have entirely neglected the friction between the screw and the nut, which is in every case far from negligable. Hence in practice P has to be considerably greater than $\frac{h}{2\pi r}W$.

95. The Balance.—As has been mentioned in § 90, the ordinary balance is essentially a lever of the first class, in which the arms are of equal length. In the balance, the "power" and "weight" consist of the force exerted by gravity on the masses carried in the pans.

The conditions which a good balance has to fulfil are as follows: (1) The balance must be true. That is, the beam must remain horizontal whenever equal masses are placed in the scale-pans. (2) The balance must be sensitive. That is, a small difference in the masses in the two pans must cause an appreciable deviation of the beam from its horizontal position. (3) The balance must be stable. That is, the beam, after being disturbed from its equilibrium (*i.e.* horizontal position), must return to it again. (4) Practically it is advisable that the period of the balance beam, when it is disturbed and oscillates before again coming to rest, should be as small as possible.

In a good balance, the knife-edges from which the scale-pans are suspended and the central knife-edge are all parallel, and lie in the same plane. If this latter condition is not fulfilled, the sensitiveness of the balance will vary with the load in the pans. For suppose ABC (Fig. 68) is a horizontal line through the central knife-edge, and that the end knife-edges are below this line at D and E. Then if each pan with its load exerts a vertical force W, and the distances between the end knife-edges and the central knife-edge are equal, the beam will be in equi-

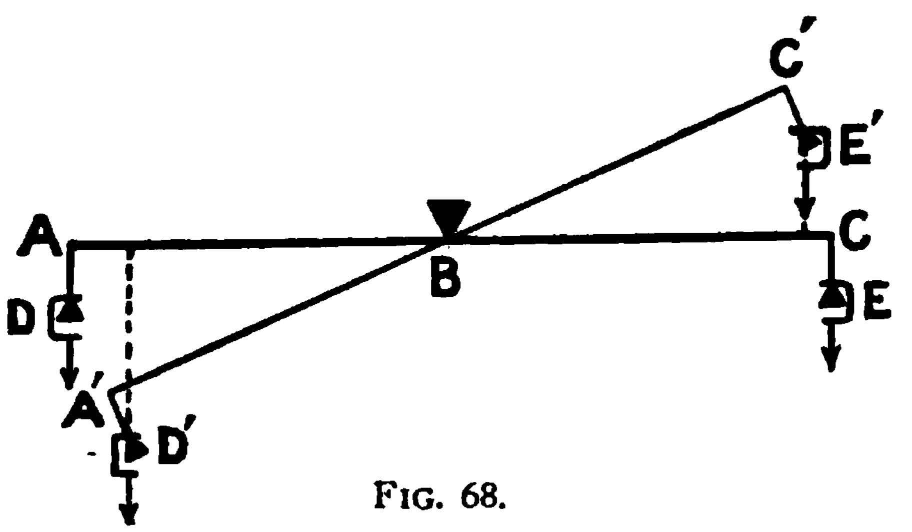


FIG. 68.

librium. If the beam is now displaced into the position A′BC′, it will be seen that the arm at which the force W at D′ acts is less than the arm of the force W at C′. Hence the moments of these two equal forces about B are not equal, and the resultant moment tending to bring the beam back into its equilibrium position is equal to W multiplied by the difference between these arms. Therefore, since this moment depends on W or the load in the pans, the greater W is, the greater will be the tendency for the beam to return to its equilibrium position, so that it will require a greater *difference* in the loads in the two pans to deflect the beam. In the following discussion we shall assume that the central knife-edge and the terminal ones all lie in a straight line.

Under these conditions, let A and C (Fig. 69) be the points of support of the scale-pans, formed by the terminal knife-edges, and B the fulcrum, formed by the central knife-edge, and let G be the point at which the weight w of the *beam* acts, *i.e.* its centre of gravity (§ 110). If S_1 and S_2 are the weights of the scale-pans, and P and W the weights of the masses placed in them, then there will be a vertical force $P+S_1$ acting at C and one of $W+S_2$ acting at A. If the beam is to remain horizontal when the scale-pans are removed, G must be vertically under B, for otherwise the force w would have a moment about B, causing the beam to be deflected. If the balance is to be true, the beam must remain horizontal whenever equal masses are placed in the scale-pans. Hence the beam must be horizontal when there are no loads in the pans, and also when each pan is loaded with a mass P. If the length of the arm $\overline{AB}$ is a and that of BC is b, then we must have in both of the above cases the moments about B equal. That is,

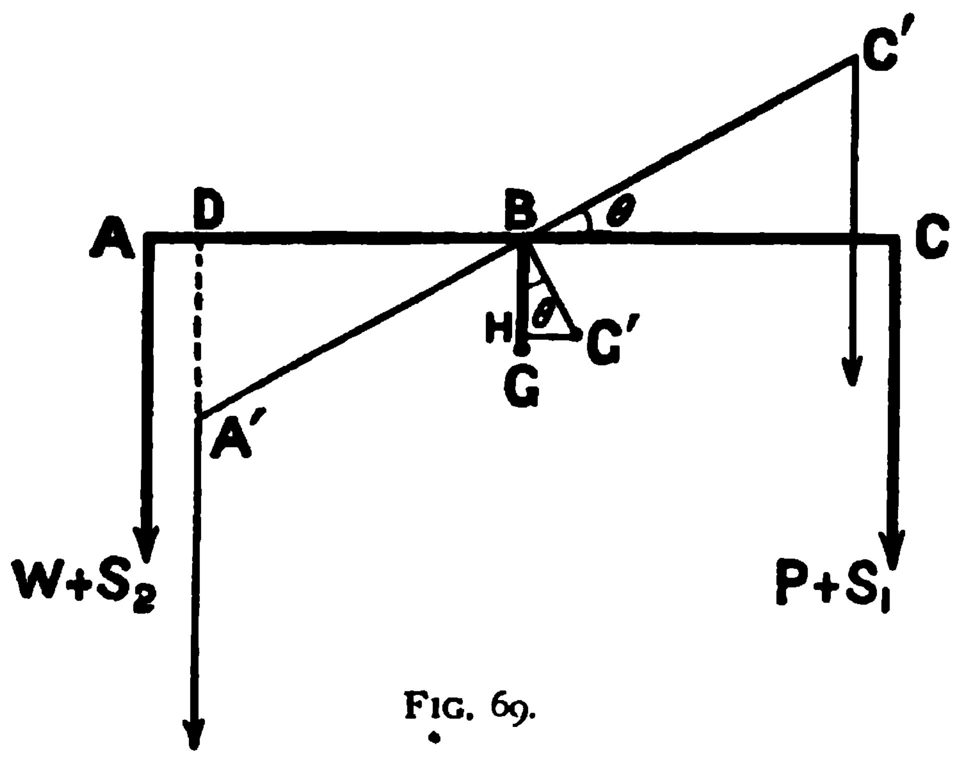


FIG. 69.

$$S_1 b = S_2 a,$$

and $$(S_1+P)b=(S_2+P)a.$$

Subtracting the first of these equations from the second, we get

$$Pb=Pa,$$

or $$b=a.$$

Also, since $$S_1 b = S_2 a,$$

$$S_1=S_2.$$

We see therefore that, for the balance to be true, the arms must be of equal length ($a=b$) and the scale-pans of equal weight ($S_1=S_2$).

In order to obtain the conditions for sensitiveness we may suppose that the load in the right pan is P and that in the left $P+x$, and that the beam takes up the position A′BC′, making an angle θ with the horizontal. The turning moments in the positive direction is then equal to $x.\overline{DB}$, while that in the negative direction is $w.\text{HG}'$, where w is the weight of the beam and $\overline{HG'}$ is the distance between its centre of gravity in the displaced position and the vertical through the central knife-edge. Since the angle GBG′ is equal to θ, we have

$$\overline{HG'}=\overline{BG'}\sin\theta=h\sin\theta,$$

where h is the distance of the centre of gravity of the beam below the central knife-edge.

Also $$\overline{DB}=A'B\cos\theta=a\cos\theta.$$

Hence $$x.a\cos\theta=w.h\sin\theta,$$

or $$\tan\theta=\frac{a}{wh}.x.$$

Now the greater the value of tan θ, the greater must be the angle θ, *i.e.* the greater the deflection. Thus for a given difference (x) in the loads of the pans the magnitude of the deflection will depend on the magnitude of the fraction $\frac{a}{wh}$. This fraction is increased in value if we increase a or decrease w or h. Therefore, in order that the balance may be as sensitive as possible, so that a readable deflection may be produced by a small difference in the masses placed in the pans, we must make a, that is the length of the beam, as large as possible, and make w, the weight of the beam, and h, the distance of the centre of gravity below the central knife-edge, as small as possible.

In order that a balance may be quick in returning to its position of equilibrium after being displaced, it is necessary that, when displaced, the moment tending to bring the beam back to its equilibrium position should be as large as possible. Since, when the pans are equally loaded, the only turning moment is that due to the weight of the beam, to secure quickness of vibration we must make the quantity $wh\sin\theta$, which expresses this moment, as large as possible for every value of θ. We can do this by making h large. No advantage would accrue by making w large, since, although we should thereby increase the turning moment, we should increase the mass to be moved in the same proportion, so that the acceleration with which the beam would move, and hence the time taken to return to its equilibrium position, would remain unaltered.[1] The

[1] The reason for this will be more apparent when the subject of the time of vibration of a pendulum has been discussed. See § 113.

only way, therefore, of securing rapidity in the indication of a balance is to make the centre of gravity of the beam some distance below the central knife-edge. It will be observed that this condition is in direct antagonism to one of the conditions for sensitiveness, and we have in this case to choose such a value for h as will make the balance fairly quick, without unduly reducing the sensitiveness. Another element which affects the quickness of a balance lies in the fact that when the beam moves the mass moved includes not only the beam itself, but also the scale-pans and their contents. It will be evident that for a given load the distance through which the load is moved, as the balance beam swings, is greater if the beam is long. Hence lengthening the beam will increase the time the balance takes to swing. This requisite, again, clashes with one of the requisites for sensitiveness.

The easiest way of drawing attention to the way in which the different requisites of a good balance are secured in a modern physical balance is to describe such a balance; and of the many slightly different types in use we will select one of those made by Bunge. The beam of the balance consists of a triangular girder-shaped framework ABC (Fig. 70). This framework carries the central knife-edge H and the end knife-edges A and B. It also carries in front a notched cross-bar DE, on which the rider can be placed, and a long pointer F. An upright rod attached to the back of the beam serves to counterbalance the pointer, &c., in front, and carries two small weights I, by means of which the position of the centre of gravity of the beam may be raised or lowered, and hence the sensitiveness altered; by moving the weight on the horizontal arm, the beam can be made to balance in the horizontal position when there is no load on the pans. The stirrups which carry the pans have small agate planes, which rest on the terminal knife-edges of the beam. The stirrups also carry two small agate points which, when the beam is lowered, fall, one into a small conical hole, and the other into a groove, which are carried by uprights K attached to the stand. These serve to slightly raise the agate planes from the knife-edges when the beam is lowered, and thus prevent the knife-edges being damaged when weights are being placed on or removed from the pans. The beam itself, when lowered, is raised from the central plane by two similar agate pins, as well as by two knife-edges LL, which support the arms. The position of the beam is read by the pointer F, which moves over an ivory scale. For very accurate work, where the smallest movement of the pointer has to be observed, a microscope M is employed, which is focussed on a small, finely divided scale G attached to the pointer. The handle N serves to raise and lower the beam, and to raise the supports which come up and catch the lower surface of the pans when the beam is lowered. A small lever O, worked by the handle P, serves to adjust the position of the rider. The rider itself weighs only half a centigram, and the position of the adjusting weights, I, is so chosen that the beam is horizontal when the rider is at

the extreme left-hand end of the beam, and no weights are in the pans. Thus when the rider is at the centre of the beam it is equivalent to a weight of half a centigram in the right-hand pan, while when it is at the extreme right-hand end of the beam it is equivalent to a centigram in this pan. The object of this arrangement, rather than the more usual one where the rider weighs a centigram and only moves over half the length of the beam, is that the scale along which the rider moves is twice as open. This is of importance, since the length of the beam is

Fig. 70.

only thirteen centimetres, so that otherwise the movement of the rider, corresponding say to a tenth of a milligram, would be so small as to be hardly observable. The advantage of the short beam is that the time the balance takes to make a swing is much smaller than would be the case with a long beam, so that the time taken to make a weighing is thereby much reduced. By the employment of a very long pointer and the microscope, we make up for the sensitiveness lost by the use of a short beam.

CHAPTER XII

FRICTION

96. Statical Friction.—Suppose that a body C (Fig. 71), of mass m, rest upon a horizontal plane AB. Then, if no force except gravity acts, C will be in equilibrium under the action of two forces—(1) the weight mg of the body acting vertically downwards, and (2) the reaction of the plane, which must act vertically upwards and be equal to mg. Now, let a force P act on C, parallel to the surface $\overline{AB}$. It is found that unless P exceeds a certain value the body still remains at rest. Under these circumstances the body is in equilibrium under its weight mg, the force P and the reaction between its surface and the plane, which must now be inclined to the normal, and act in some such direction as CR. This force along CR′ may be resolved into a reaction normal to the surface, *i.e.* along CR, and a force along CF which must, if there is equilibrium, be equal in magnitude to P. This force, which is brought into play when we attempt to slide one body over another, and which *always* acts so as to *resist* motion, is called the *friction* between the surfaces.

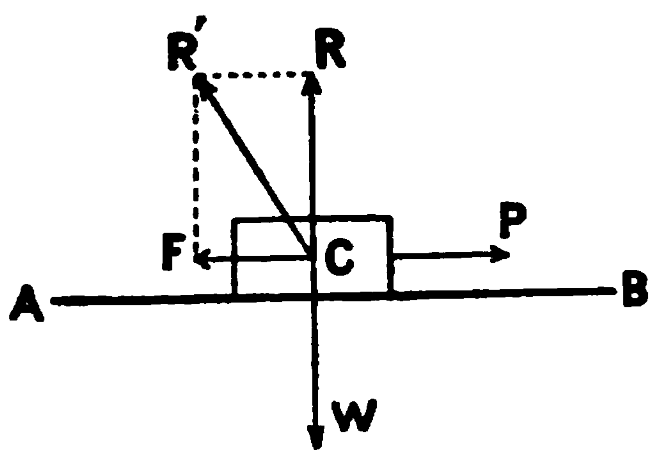


FIG. 71.

If the total *normal* pressure between C and the plane be Q, then it is found that C will commence to slide when the force P bears to Q a certain ratio, which is necessarily less than unity. This ratio is called the *coefficient of friction* between the body C and the plane AB, and is generally denoted by the symbol μ. The value of the coefficient of friction is independent of the size of the surface of contact between C and AB, and of the pressure Q. It depends, however, on the nature of the substances forming the two surfaces in contact, on the smoothness of these surfaces, and on the presence or absence of any lubricant, such as oil, fat, blacklead, &c., between the surfaces. The value of μ has to be determined experimentally for each of these conditions.

If the force P is less than μQ, then there will be no motion, and the frictional resistance F will be equal and opposite to P. When P is just equal to μQ motion will be on the point of taking place, and the frictional

resistance will have its maximum value (μQ). If P is greater than μQ motion will take place, but the moving force will be less than P, since, although when motion has commenced the frictional resistance is often no longer equal to μQ, yet friction still acts as a force tending to prevent motion.

Since the *coefficient* of friction is independent of the surface of contact, it follows that for a given value of Q the frictional resistance (F) is also independent of the extent of the surface of contact. If A is the area of this surface, then the pressure per unit area is Q/A, and the frictional resistance per unit area is μQ/A. If, while Q remains the same, A is reduced to A′, then the pressure per unit area is increased to Q/A′, and the frictional resistance per unit area is increased to μQ/A′. Hence the frictional resistance *per unit area* varies directly as the pressure per unit area.

97*. Limiting Angle.—When motion is just about to commence, and hence P is equal to μQ, the body is in a state of equilibrium under three forces, the force P acting horizontally, the pressure Q acting vertically downwards, and the reaction acting along CR′ (Fig. 72). In order to find the angle which CR′ makes with the normal, we draw a line DE parallel to Q, and of such a length that it represents Q in magnitude, and from E draw $\overrightarrow{EF}$ parallel to P, and hence at right angles to $\overline{DE}$, to represent P in magnitude. Then, by the triangle of forces (§ 72), the reaction which, together with the forces P and Q, maintains the body C in equilibrium, must be represented in magnitude and direction by the line $\overrightarrow{FD}$. Therefore the angle FDE is equal to the angle ϕ between the reaction CR′ and the normal. Since $\overrightarrow{DE}$ is equal to Q and $\overrightarrow{EF}$ to P, which is equal to μQ, we have

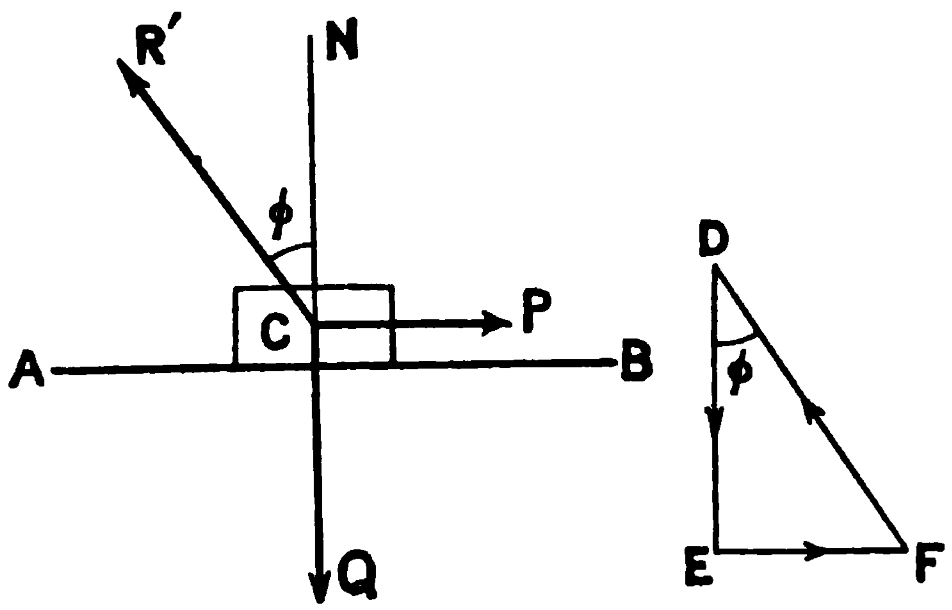


FIG. 72.

$$\tan \phi = \frac{\mu Q}{Q} = \mu.$$

This angle ϕ, which represents the greatest angle the line of action of the reaction can make with the normal to the surface of contact, is called the *limiting angle*.

If a force is applied to C along such a direction as $\overrightarrow{FC}$ (Fig. 73), making an angle of ψ with the normal, then if ψ is less than the limiting angle, motion of C will not take place, however great the value of this force. The reason is that we may resolve the force into two components, one parallel to the surface, which tends to produce motion and is resisted by the friction, and the other, which acts along the normal, produces a contact pressure. If F is the force, the component parallel to the surface is F sin ψ, and the component parallel to the normal is F cos ψ. If motion is just about to take place, and we neglect the weight of the body, then

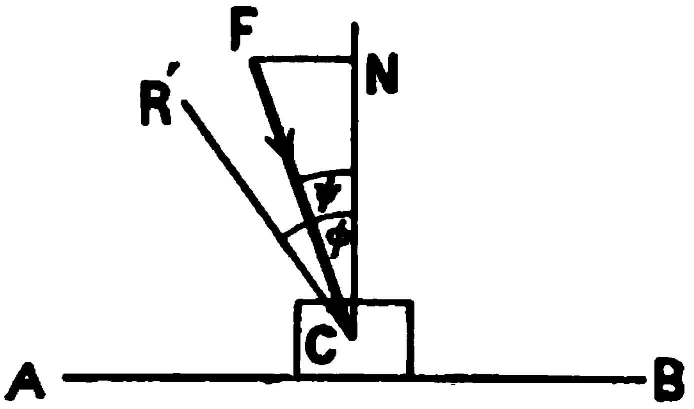


FIG. 73.

$$F \sin \psi = \mu F \cos \psi,$$
$$\therefore \quad \mu = \tan \psi.$$

But $\mu = \tan \phi$ where ϕ is the limiting angle. Hence if ψ is less than ϕ motion will not take place.

98*. Angle of Repose.—If a body G (Fig. 74) of mass m is placed on an inclined plane AB, then, if there were no friction between G and the plane, the only forces acting would be the weight, which is a force of mg acting vertically downwards and the reaction of the plane $\overrightarrow{GR}$ acting at right angles to AB. As these forces are not in the same straight line, the body would move down the incline. If, however, there is friction between G and the surface of the plane, the friction will tend to prevent motion, and till the plane has a certain slope the body will remain at rest. To find the maximum inclination (ψ) of the plane to the horizontal we resolve the force mg into a component parallel to BA, which tends to produce motion, and a component normal to BA, which acts as the contact pressure. In the triangle DGE, the angle EGD is equal to ψ, and ED is parallel to AB. Hence the component of mg parallel to BA is $mg \sin \psi$, and the component perpendicular to BA is $mg \cos \psi$. If motion is just about to commence,

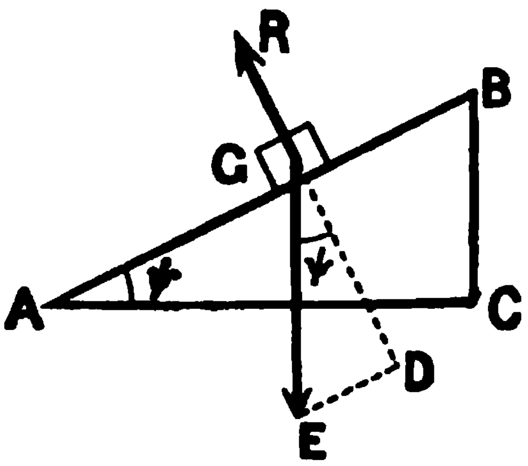


FIG. 74.

$$mg \sin \psi = \mu mg \cos \psi$$
$$\therefore \quad \mu = \tan \psi.$$

Hence if ψ is greater than the limiting angle, motion takes place. The maximum inclination to the horizontal of the plane which is possible without the body sliding is called the *angle of repose*.

Thus the angle of repose is equal to the limiting angle, and the coefficient of friction is equal to the tangent of either of these angles.

99*. Kinetic Friction between Solids.—As mentioned in § 96, after slipping has commenced the friction continues as a force tending to prevent motion, but the magnitude of the friction is in general less than it is just before slipping commences. It is found by experiment that, as long as the speed of the motion is not too great, the frictional resistance is proportional to the total pressure between the two solids, and independent of the velocity.

If Q is the total normal pressure between the solids and F is the frictional resistance, then

$$F = \nu Q,$$

where ν is called the coefficient of *kinetic friction*. Hence if a force P parallel to the plane surface AB (Fig. 72) act on a body C of mass m, then $Q = mg$ and $F = \nu mg$. Since the frictional resistance opposes the motion, the resultant force which is available for changing the motion of the body is $P - F$ or $P - \nu mg$. The acceleration (a) produced by this force is given by

$$a = \frac{P-F}{m} = \frac{P}{m} - \nu g.$$

If there had been no friction the acceleration would have been P/m, so that the effect of friction is equivalent to a negative acceleration of νg units.

Of course, if P is less than νmg, the body if in movement will gradually come to rest, and then it will require a force greater than μmg to start motion again. In the following table, some of the values of the coefficient of kinetic friction obtained by Morin are given.

COEFFICIENT OF KINETIC FRICTION.

Oak on oak . .	Fibres parallel .	Surfaces without lubricant .	0.48
" "	" "	Surfaces rubbed with dry soap	0.16
" "	" perpendicular	Surfaces without lubricant .	0.34
" "	" "	Surfaces wetted with water .	0.25
Iron on oak . .	Fibres parallel to movement	Surfaces without lubricant .	0.62
" "	" "	Surfaces rubbed with dry soap	0.21
Iron on bronze	...	Surfaces slightly greasy . .	0.18
Iron on iron, steel on steel, oak on oak, bronze on bronze . .	...	Surfaces slightly greasy . .	0.15
		Surfaces greasy	0.07–0.08
		Surfaces greasy and lubricant continually renewed . .	0.04–0.05

This table shows in a very marked manner how the friction between solids is reduced by the presence of a layer of lubricant. If the motion is extremely rapid, and the surfaces in contact are sufficiently large, it is possible to use air as a lubricant, and under these circumstances the friction is enormously reduced.

100*. Rolling Friction.—When a wheel or cylinder rolls on a plane surface, there is produced at the point of contact a resistance to the motion which is generally said to be due to *rolling friction*. This resistance is not a true friction in the sense of the word used in previous pages, since there is no relative motion of the two surfaces at the points of contact, hence there is no slipping.

Suppose that a cylinder EF (Fig. 75) rolls on a horizontal plane AB, and a light string is passed over the cylinder, the tensions P and Q in the two portions of this string being adjusted so that the cylinder, when started, continues to move with a uniform speed, rolling along from A towards B. Since the motion is *uniform*, it follows that the forces acting on the cylinder are in equilibrium. These forces are the weight W of the cylinder acting vertically downwards through the axis G, the forces P and Q, and the reaction of the plane AB. Now it is found experimentally that, if the motion in the direction from A to B is to be uniform, Q must be greater than P. The resultant of the parallel forces Q and P will therefore be a force nearer Q than P (§ 69). Let $\overrightarrow{HK}$ be this resultant. The resultant of this force HK and the weight of the cylinder must lie somewhere between them, say along $\overrightarrow{LM}$. If then the forces are in equilibrium, the reaction of the plane must act along $\overrightarrow{LN}$. In other words the reaction of the plane does not act through the point C, where in the figure the cylinder touches the plane. This apparent impossibility is explainable if we suppose that rolling friction is really due to the fact that the plane becomes deformed and a small ridge is "rolled" up in front of the cylinder, or that the cylinder itself becomes flattened. The former of these effects can be clearly seen if a wheel is rolled on a sheet of india-rubber; for, as shown in Fig. 76, the rubber is forced up into a small ridge before the wheel. The latter effect is illustrated in the case of a pneumatic bicycle tyre. The magnitude of the resistance to motion in the case of rolling is very much smaller than

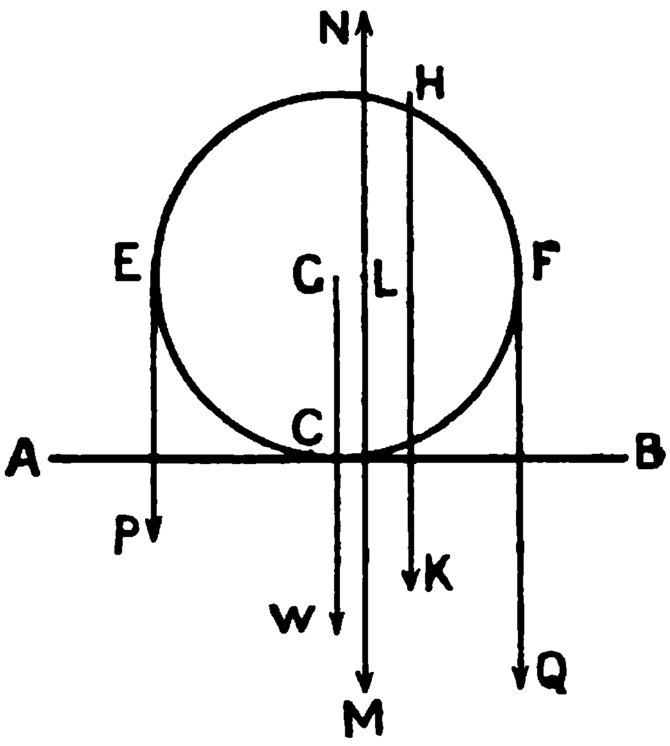


FIG. 75.

that in the case of sliding. Thus Coulomb found that in the case of a cylinder of lignum-vitæ, 16 centimetres in diameter, when loaded with 1000 pounds, the resistance to rolling amounted to 6 pounds, while with the same load the resistance to sliding would have amounted to at least 200 pounds. Whenever it is possible, it is therefore advantageous to substitute rolling for sliding, if the frictional resistance to motion is to be

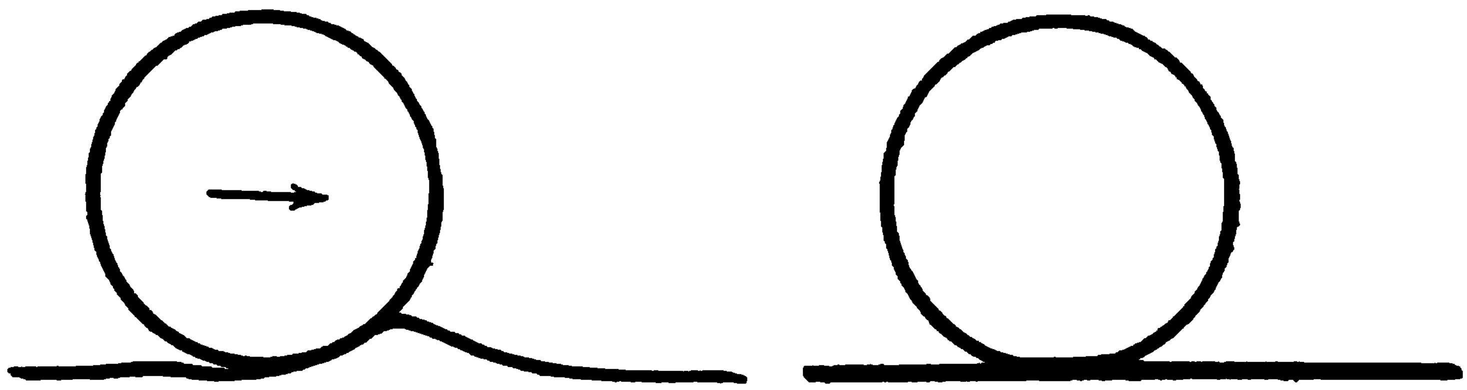


FIG. 76.

reduced. Thus, in the case of a carriage, the sliding friction which occurs in a sledge is replaced by rolling friction between the tyre of the wheel and the ground. In the modern bicycle even the sliding friction of a wheel upon its axle is, as far as possible, replaced by rolling friction in the ball-bearing, where a number of hard steel balls are placed, so that the hub of the wheel rolls on them, and they roll on the axle itself.

101. Loss of Available Energy due to Friction.—Since in every case friction acts as a force tending to check the motion, whenever any displacement actually takes place work will have to be done against the frictional resistance. The energy which is necessary to perform this work is converted into heat, and this heat gradually becomes diffused amongst neighbouring bodies, and so the energy is no longer available for doing work. The frictional resistance *always* opposes motion, so that if we change the direction of motion the direction of the frictional resistance also changes, so that it is impossible to utilise this force to increase the motion of a body or to do work, but work has always to be done *against* it. It is therefore hardly correct, in view of the definition of force given in § 59, to call the frictional resistance a *force*. Since, however, it always acts as if it were a force opposing motion, it is convenient so to regard it.

102. Friction-Dynamometer.—One of the applications of friction is to employ it to measure the power or rate of doing work of a machine, such as a steam-engine. A form of friction-dynamometer for this purpose is shown in Fig. 77. A pulley A with a flat edge is fixed to the shaft of the engine, and a strap BCD, on the inside of which blocks of wood are usually fixed, rests on the edge of this pulley. One end of the strap is attached to a spring balance E, by means of which the tension acting on this end of the strap can be measured, while a tension P, caused by a weight suspended on the other end, serves to keep the strap tight. The

engine having been started, P is increased till it is exerting its maximum power; the work being done against the friction of the wooden blocks on the edge of the pulley. If r is the radius of the pulley and R the radius of the portion of a circle along which the strap lies, n the number of turns the pulley makes per second, and the reading on the spring-balance is W. Then the sum of the moments of P and W about the axle is $(W-P)R$, and this must be equal to the moment of the friction about the same axis, since the strap and blocks are in equilibrium. Hence, if F is the frictional resistance,

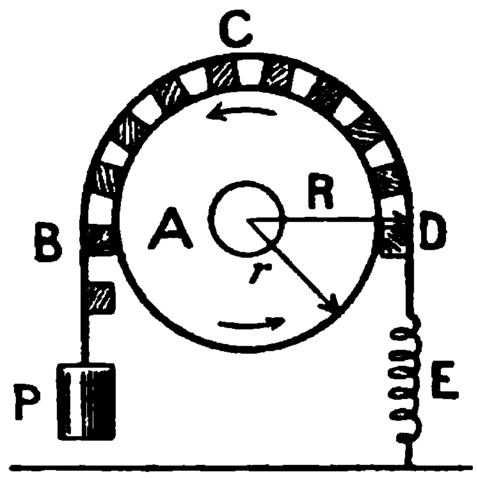


FIG. 77.

$$Fr=(W-P)R \quad . \quad . \quad . \quad (i).$$

Now the frictional resistance F acts tangentially to the pulley, and tends to check the motion. The distance through which the edge of the pulley moves against F during one second is $2\pi rn$. Hence the work done against friction in one second is

$$2\pi rn.F \quad . \quad . \quad . \quad . \quad (ii).$$

and this is the power expended in friction. Substituting the value of F obtained from equation (i) in (ii), we get that the power spent against friction is

$$2\pi rn.\frac{(W-P)R}{r}$$

or

$$2\pi nR\,(W-P).$$

If then the whole available power of the engine is spent in overcoming the friction of the dynamometer, and H is the number of units of work per second, in the system in which R, P, and W are measured, which are equal to a horse-power, the horse-power will be

$$\text{HP}=\frac{2\pi nR(W-P)}{H},$$

in which n is obtained by counting, R by adding the thickness of the wooden blocks to the radius of the pulley, and P and W are obtained from the weights placed in the scale-pan and the reading of the spring balance respectively.

CHAPTER XIII

GRAVITATION

103. Attraction and Repulsion.—In the case of two portions of matter between which a stress exists, and in which we are unable to trace any material connecting link, it is usual, if the stress tends to make the bodies move towards one another, to say that the bodies *attract* one another. If, on the other hand, the stress is such as to tend to make the bodies separate, then we say they *repel* one another. For convenience, the force in the case where the bodies *repel* one another is generally regarded as *positive*, and that in the case where they *attract* one another as *negative*.

104. The Law of Inverse Squares.—In general, when the distance between two particles which attract or repel one another changes, the stress between the particles alters. The only case which we shall examine, since it is by far the most important in physics, is that in which the stress varies inversely as the square of the distance between the particles, and takes place in the direction of the line joining them. As a particular case of this general law, which is called the law of the inverse square, we may take the case of the attraction between two particles of mass m and m'. The stress between the particles depends directly on the product of the masses and inversely on the square of the distance (d) between them. Hence if F is the force which either particle exerts on the other,

$$-F \propto \frac{mm'}{d^2},$$

or

$$-F = k.\frac{mm'}{d^2},$$

where k is some constant.

If, instead of having only two particles, there are a number, then experiment shows that the force exerted on each particle is the resultant of all the forces which would be exerted by each of the other particles separately.

105*. Work done by Attraction or Repulsion.—If the distance between two particles which attract one another is increased, work will have to be done on the system. If the particles approach one another, however, they are capable of doing work. The maximum of work will be done by or on the system when the particles are brought from an infinite

distance into contact, or moved from contact to an infinite distance from one another.

It therefore follows that, in the case of two particles which attract one another, when they are not in contact they possess a certain amount of potential energy, for they can be made to do work during their approach. The amount of work which can be done when the system changes from the positions shown at A, B (Fig. 78) to the positions A, B′ depends simply on the positions of B and B′ with reference to A, and is independent of the form of the path by which the particle passes from B to B′. If this were not so, suppose that when the particle went from B to B′ along the path BCB′ more work was done by the particle, than when it went along the path BDB′. Now suppose the work done by the particle when taken along the path BCB′ is w_1, and that along the path BCB′ is w_2, so that $w_1 > w_2$. Let the particle move from B to B′ along the path BCB′, doing w_1 units of work, and be taken back to B along the path B′DB, w_2 units of work being done on it. The configuration of the system is now the same as at first, but in the cycle performed by the particle B the work w_1 done by the system is greater than the work w_2 done on the system, hence there remains a quantity of energy $w_1 - w_2$ over, which might be utilised for doing external work, so that the arrangement would give us "perpetual motion." Since by the doctrine of the conservation of energy this is impossible, it follows that w_1 must be equal to w_2, or, in other words, the work done either by or on the system between any two configurations depends solely on the initial and final configuration (*i.e.* the positions of the particle), and is independent of the intermediate steps by which the system changes from one configuration to the other.

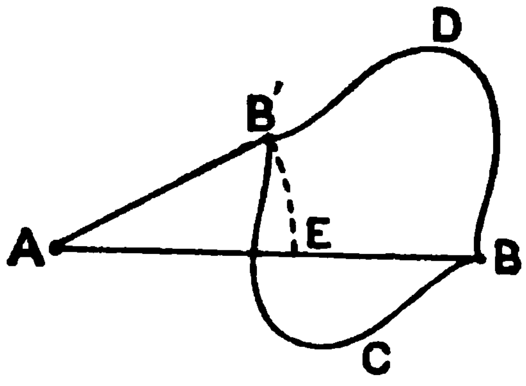


FIG. 78.

In order to calculate the work which can be done by the particle in passing from B to B′, we suppose the path to consist of two portions, first a part $\overline{BE}$ of the straight line joining B to A, and then a portion of a circle with A as its centre and $\overline{AB'}$ as its radius. Since there will be no work done during the passage of the particle along the circular portion of the path, the motion being at all points at *right angles* to the force, the work done along the path $\overline{BE}$ will be equal to the work done along any path between B and B′.

When the particle is at B, the force acting on it is $\frac{kmm'}{d_1^2}$; if it moves through a small distance δd, the force becomes $\frac{kmm'}{(d_1 - \delta d)^2}$. The product of the mean force into δd gives the work done over the small element δd of the path. The work done over the next element δd is in the same way

$\frac{kmm'}{2}\left[\frac{1}{(d_1-\delta d)^2}+\frac{1}{(d_1-2\delta d)^2}\right]\delta d$, and so on. The total work done will be the sum of a number of such terms taken for the whole path.

If d_2 is the distance $\overline{AB'}$ or $\overline{AE}$, then the work done between B and E, and therefore also between B and B′, can be shown[1] to be $kmm'\left(\frac{1}{d_2}-\frac{1}{d_1}\right)$.

If the point B is at an infinite distance from A, then the work done by the body in the case of attraction, or against the body in the case of repulsion, when moved up from an infinite distance to a distance d_2 from A, is

$$\frac{kmm'}{d_2},$$

since $\frac{1}{\infty}=0$.

If the body which is moved is of unit mass, *i.e.* if $m'=1$, then the work done is

$$\frac{km}{d_2}.$$

In the case of two bodies which repel one another, and are at a distance d_2, they possess a certain potential energy due to their mutual repulsion. The amount of this potential energy is $\frac{kmm'}{d_2}$, since this expression gives the maximum amount of work which could be done by the mutual repulsion of the bodies, for there is no force exerted between the bodies when they are at an infinite distance, so that their potential energy, as far as their mutual repulsion is concerned, is zero.

106*. Potential.—It has been shown, in the last section, that the work which has to be done to remove a unit mass from a given point in the neighbourhood of another mass, to a place where there ceases to be any attractive or repulsive force between the masses, is a fixed quantity depending simply on the position of the point with reference to the attracting mass. This quantity of work may be looked upon as an attribute of the given point, the attracting mass of course being supposed to remain fixed in position, and it is then called the *potential* of the given point. Thus the potential at a point at a distance d from a mass m is equal to

$$\frac{km}{d}.$$

107. Kepler's Laws.—Astronomical observations having shown that the earth and the other planets move round the sun in approximately circular (really elliptical) orbits, it follows that there must be attraction between each planet and the sun, for otherwise the planets would travel in straight lines.

Kepler, by a careful study of the observations on the motion of the planets made by Tycho Brahe, deduced three laws which now bear his name.

[1] See § 464.

Kepler's laws are as follows :—

1. The areas swept out by the straight line joining a planet to the sun are proportional to the time. Thus in Fig. 79, if S is the position of the sun, and $P_1P_2P_3 \ldots P_6$ is the orbit of a planet, and in a given interval t the planet moves from P_1 to P_2, or from P_2 to P_3, or from P_3 to P_4, &c., then the areas P_1SP_2, P_2SP_3, and P_3SP_4, &c., are all equal.

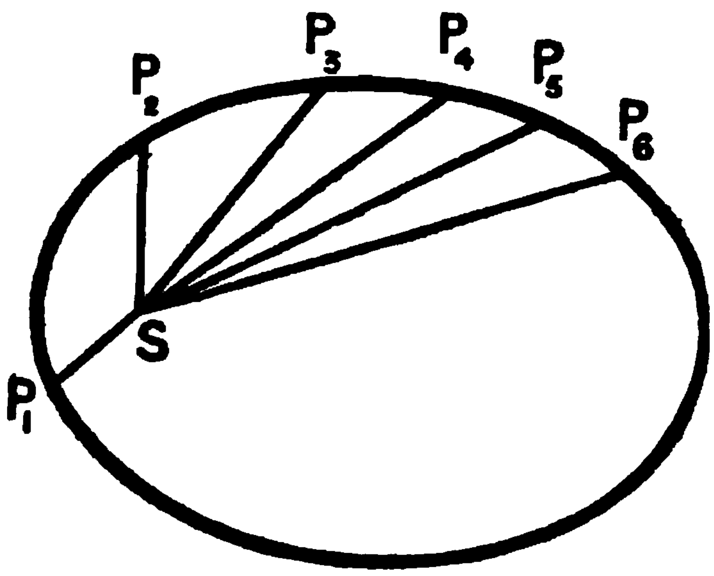


FIG. 79.

2. The orbit of a planet is an ellipse, the sun being at one of the foci.

3. The squares of the time taken to describe its orbit by different planets are proportional to the cubes of the mean distances of the planets from the sun. Thus if T_1 and T_2 are the times taken by two planets to describe their orbits, and D_1 and D_2 are their distances from the sun, then Kepler's third law states that

$$\frac{T_1^2}{T_2^2}=\frac{D_1^3}{D_2^3}.$$

108. Newton's Law of Gravitation.—Although Kepler's laws give us a description of the motion of the planets, they do not tell us anything about the *forces* which serve to determine these motions. Newton, however, discovered the dynamical interpretation of Kepler's laws, and showed that if we suppose that a stress is set up between each of the planets and the sun directly proportional to the mass of the planet and inversely proportional to the square of the distance of the planet from the sun, then the motions of the planets will be just such as would satisfy Kepler's laws. Although it had been previously suggested that the sun as a whole attracted each planet as a whole, and the law of the inverse square had even been enunciated, it is to Newton that we owe the law of gravitation in the form in which it remains to this day, viz. every portion of matter attracts every other portion of matter, and the stress between them is proportional to the product of their masses divided by the square of their distance apart.

As a test of the truth of his law, Newton showed that it correctly accounted for the force necessary to retain the moon in her orbit.

If we assume that the orbit of the moon (with reference to the earth) is a circle of radius R, then by § 42 the acceleration of the moon towards the earth necessary to keep it moving in this orbit will be

$$a=\frac{v^2}{R},$$

where v is the linear velocity with which the moon is moving in the

circular orbit. If T is the time the moon takes to complete the orbit (a sidereal month), then

$$v = \frac{2\pi R}{T}.$$

Hence

$$a = \frac{4\pi^2 R}{T^2}.$$

Since the force exerted by the attraction of the earth on a given mass is proportional to the acceleration produced in the mass (§ 61), it follows that if Newton's law is true, *i.e.* if the force decreases as the square of the distance, then if r is the radius of the earth and g the acceleration produced by gravity at the surface of the earth, the acceleration (a') produced by the gravitational attraction of the earth at the distance of the moon will be given by

$$\frac{a^1}{g} = \frac{r^2}{R^2}$$

or

$$a^1 = \frac{gr^2}{R^2}.$$

In order to calculate the values of a and a^1 we may take

$R = 240{,}000$ miles.
$r =$ 4000 miles.
$T =$ 27.3 days.
$g =$ 32.2 feet per sec. per sec.

Since g is expressed in feet per second per second, we must reduce R and r to feet and T to seconds, then substituting we get

$$a = \frac{4\pi^2.240000.5280}{(27.3.86400)^2} = .00899 \text{ ft./sec}^2.$$

$$a^1 = \frac{32.2(4000)^2}{(240000)^2} = .00894 \text{ ft./sec}^2.$$

The agreement is as good as the approximate values we have assumed for the various quantities will allow.

109. The Cavendish Experiment.—The calculation given in the last section shows that the moon is attracted by the earth with a force which follows the same law as the attraction exerted by the earth on bodies on its surface. We now proceed to show that two bodies of such a size that we can handle them attract *one another*. The experimental difficulties of carrying out this investigation are, however, very great, since the mass of the largest body which we can employ is so excessively small as compared with the mass of the earth, and hence the attraction between any two bodies we can use is only a small fraction of the weight of either.

The first apparatus for measuring the gravitational attraction of two bodies was designed by the Rev. John Michell, but he died before he

had time to try the experiment. Michell's apparatus, after his death, came into the possession of Henry Cavendish, who, after slightly altering it, carried out the experiment which has subsequently been known as the Cavendish Experiment.

The instrument employed is called a torsion balance,[1] and consists of a horizontal rod suspended by a long fine wire. If this rod is acted upon by a couple in the horizontal plane, it will turn, the wire becoming twisted. Since the wire is elastic, it resists the strain (twist), tending to untwist itself. The force with which the wire tends to untwist is, within certain limits, proportional to the angle through which it is twisted. Hence, by observing the angles through which two given couples twist the wire, we have a means of comparing these couples.

The couple necessary to twist the wire through an angle θ is for a wire of any given material, say silver, inversely proportional to the length of the wire, and directly proportional to the *fourth* power of the radius. Hence, in order to get an appreciable twist when the couple is small, it is of the greatest importance that the wire should be of small diameter, and also, but to a smaller degree, a long one.

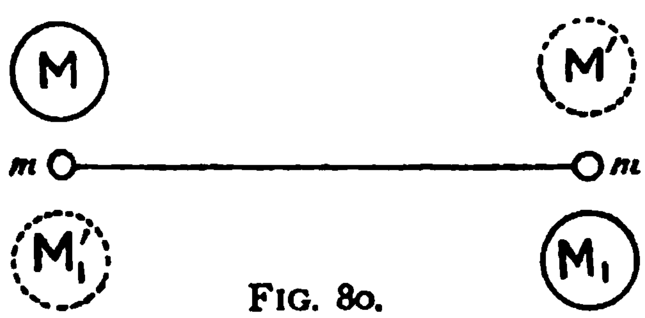


FIG. 80.

In the Cavendish apparatus two small equal masses m, m' (Fig. 80) are attached to the end of the horizontal rod of the torsion balance. Two large spheres of lead M, M_1 are supported so that they can be placed either in the positions MM_1, or $M'M_1'$. In the first of these positions, the attractions between the fixed mass M and the movable mass m, and between M_1 and m', tend to turn the beam in the direction of the motion of the hands of a clock. When the fixed masses are in the positions $M'M_1'$, however, the attraction between them and the suspended masses tends to turn the beam in an anti-clockwise direction. The position of the beam is obtained by observing, by means of a telescope, the graduations of a scale as seen reflected in a small vertical mirror attached to the centre of the beam.

If the mass of each of the suspended spheres is m and that of each of the fixed spheres M, and if d is the distance between the centre of the fixed sphere and that of the adjacent suspended sphere, then the attraction between them is $k\frac{mM}{d^2}$.

Hence, if a is the distance between the two suspended masses, the couple tending to twist the wire due to one pair of spheres is $k\frac{mM}{d^2} \times \frac{a}{2}$, or for both pairs $ka.\frac{mM}{d^2}$. If this couple produces a twist of $\theta°$ in the wire,

[1] It was subsequently reinvented by Coulomb, and is often known as Coulomb's balance.

while previous experiments have shown that the angle of twist produced by a unit couple is a, then

$$\frac{\theta}{a}=\frac{kamM}{d^2},$$

or

$$k=\frac{d^2}{amM}\cdot\frac{\theta}{a}.$$

Hence the value of k can be calculated from the observed quantities.

Since the force of attraction between a mass of one gram when at the surface of the earth and the earth is g dynes, we have, if M' is the mass of the earth, and R the radius of the earth,

$$g=k\frac{M'}{R^2}.$$

Hence

$$M'=\frac{gR^2}{k}.$$

Knowing the value of k from Cavendish's experiment, and the value of R from measurements made on the surface of the earth, we can therefore calculate the value of M', the mass of the earth. For this reason Cavendish is often said to have first weighed the earth.

Knowing the mass of the earth, then from Kepler's laws we can calculate the masses of the other planets and of the sun.

The Cavendish experiment has since been repeated by several observers, the latest measurements being those of Boys. This observer, by using an excessively fine thread of fused quartz as the suspension of the torsion balance, was able to employ comparatively small masses, the suspended spheres only weighing one gram. The results he obtained gave the value $k=6.6579\times 10^{-8}$ in *c.g.s.* units, so that two small spheres, each having a mass of one gram, when placed so that the distance between their centres is one centimetre, attract each other with a force of 6.6579×10^{-8} dynes. This value of k gives the value 5.5268 as the mean density (§ 129) of the earth.

110. Centre of Gravity.—Suppose we have two particles at A and B (Fig. 81) of mass m_1 and m_2. Then they are each attracted towards the centre of the earth, and since the centre of the earth is at a very great distance compared with the distance $\overline{AB}$, the forces exerted by gravity on the two particles will be sensibly parallel in direction. Hence we have the system consisting of the two particles acted upon by two parallel forces of magnitude m_1g and m_2g respectively. If we divide the line $\overline{AB}$ at C, so that

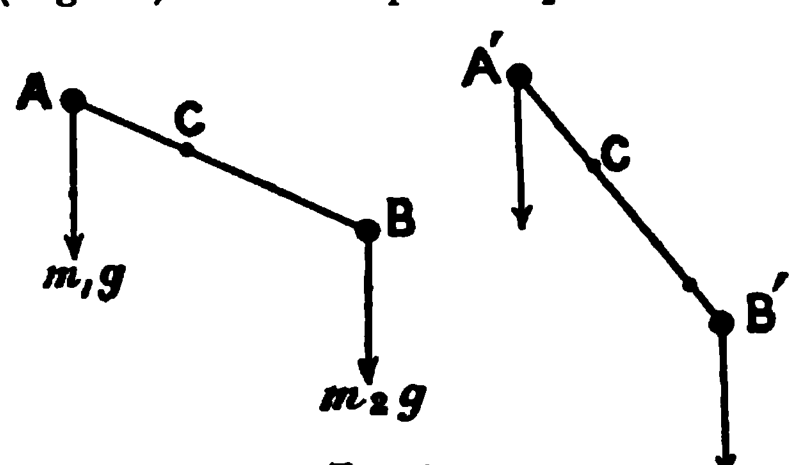


FIG. 81.

$$m_1g\times\overline{AC}=m_2g\times\overline{BC}$$

then the resultant of the two forces $m_1 g$ and $m_2 g$ will pass through C (see § 69). If the two particles are turned into any other position A′, B′, the distance $\overline{AB}$ between them being kept the same, the forces due to gravity will be inclined at a different angle to the line $\overline{AB}$, but the resultant will still pass through the point C. Hence, whatever the position of the two particles A and B, so long as their distance apart remains the same, the resultant of the gravitational attraction of the earth on the two particles always passes through a point C, which has a fixed position relatively to A and B.

If we have a system consisting of three particles A, B, and D (Fig. 82) of mass m_1, m_2, and m_3 respectively, then the resultant of the earth's attraction on A and B acts at C, as in the previous case. We may now consider that we have *two* parallel forces, one of magnitude $m_1 g + m_2 g$ acting at C, and the other of magnitude $m_3 g$ acting at D. Hence the resultant will pass through a point E, such that

$$(m_1 g + m_2 g)\overline{CE} = m_3 g \times \overline{ED}.$$

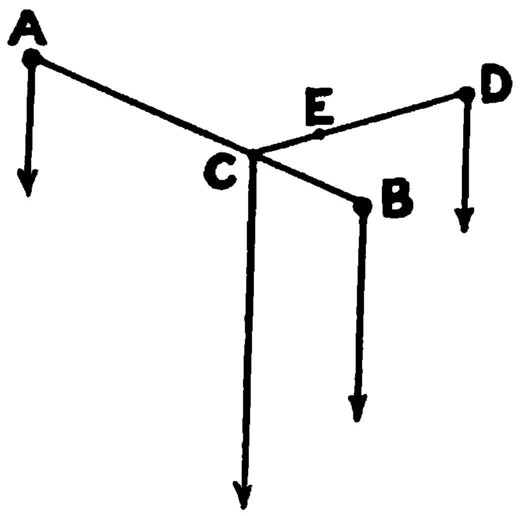


FIG. 82.

Therefore the resultant of the earth's attraction on the three particles passes through E, and will still pass through E, however the three particles are turned, so long as their *relative* positions remain unaltered.

Proceeding in this way we might find, for a system consisting of any number of particles, a single point through which the resultant of all the forces exerted by gravity on the particles will pass, whatever the position of the system. As we may regard any solid body as built up of a number of particles, it therefore follows that it will be possible in the case of every solid body to find a point, and only one point, through which the resultant of all the forces exerted by gravity on the particles constituting the body must pass. This point is called the *centre of gravity*, or the *centre of mass* of the body. From this definition it follows that the weight of a body, which is simply the magnitude of the resultant of the forces exerted by gravity on the particles which constitute the body, always acts downwards in a vertical direction, passing through the centre of gravity.

There cannot possibly be two centres of gravity belonging to one body, for if there were two, say G_1 and G_2, and the body was turned so that the line joining G_1 and G_2 was horizontal, then by definition the resultant of all the parallel forces due to gravity acting on the particles of the body passes through G_1; it also passes through a second point G_2, which is not on the line of the other resultant, which is impossible. Hence there can be only one centre of gravity.

The centre of gravity is a mathematical point, and it need not necessarily lie within the substance of the body. Thus the centre of gravity of a uniform ring lies outside the material of the ring.

If a body is suspended from a point, so that it can turn freely about the point of suspension in every direction, then the centre of gravity will lie in the vertical drawn through the point of support. Because under the circumstances given there are only two forces acting on the body, namely its weight, acting vertically downwards through the centre of gravity, and the reaction of the support. In order that the body may be in equilibrium these forces must be equal in magnitude, and act in opposite directions along the same straight line. Hence the line of action of the weight, *i.e.* the vertical through the centre of gravity, must pass through the point of support, since the reaction of the support must necessarily act through the point of support.

111. Stable, Unstable, and Neutral Equilibrium.—In the case of a body in equilibrium, when suspended from a point about which it can turn freely, two cases may occur. In each (see Fig. 83), the vertical through the centre of gravity G passes through the point of support A. In the one case, however, the centre of gravity is vertically below the point of support, while in the other case it is vertically above. In the first case, if we suppose the body slightly displaced from its position of equilibrium, as at AG (Fig. 84), we see that the weight of the body acting through G has a moment about the point of suspension A, and that this moment tends to bring the body *back* to its original position. The body is therefore said to be in *stable equilibrium*. In the other case, where the centre of gravity is above the point of support, if the body is displaced to A′G′ (Fig. 84), the turning moment due to the weight now acts so as to *increase* the displacement, and therefore the body moves further and further away from its original position. The equilibrium in which the centre of gravity is vertically over the point of support is said to be *unstable*.

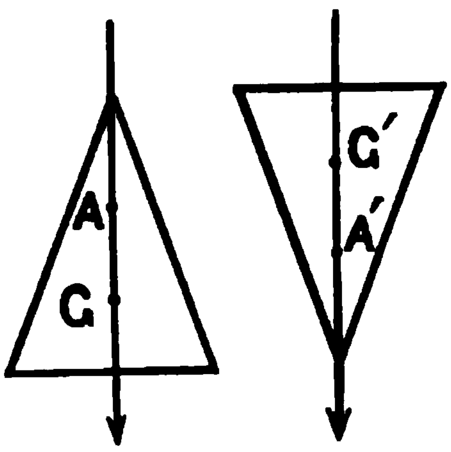


FIG. 83.

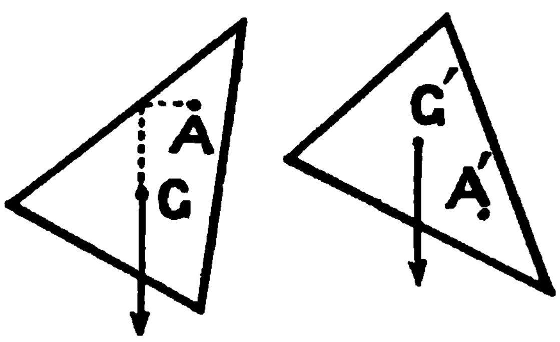


FIG. 84.

If the body is supported at its centre of gravity, then, since the weight acts through the centre of gravity whatever the position of the body, the direction in which the weight acts always passes through the point of support, and hence the body when displaced is still in equilibrium. In such a case the body is said to be in *neutral equilibrium*.

If a body is in stable equilibrium as at AG (Fig. 83), its centre of

gravity is in the lowest possible position. Its potential energy is therefore a minimum, for any possible displacement involves the raising of the centre of gravity, and in order to raise the centre of gravity of a body we must do work, because we are moving the point of application (centre of gravity) of a force (the weight of the body) against this force. Hence when a body is in stable equilibrium the potential energy is a minimum, and any possible displacement of the body involves the raising of the centre of gravity, and therefore the expenditure of external energy.

As an example of stable equilibrium we may take a book lying on a horizontal table (Fig. 85). In this position the centre of gravity of the book is as low down as possible, and any displacement, except sliding along the table which does not alter the body's state of equilibrium, such as that shown at B, involves the raising of the centre of gravity.

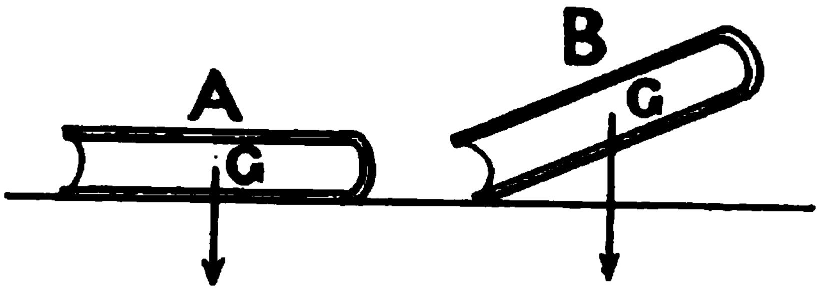


FIG. 85.

In unstable equilibrium any slight displacement tends to lower the centre of gravity, the potential energy being, in this form of equilibrium, a maximum. The potential energy tending to change into kinetic energy, a very slight disturbance may, in unstable equilibrium, produce a very great displacement, for this displacement does not require the supply of external energy to the system in order for it to take place.

In neutral equilibrium, there is no change in the potential energy when the body is displaced, the centre of gravity remaining at the same height. A sphere resting on a horizontal table is an example of neutral equilibrium.

We may somewhat generalise the above statements with reference to the equilibrium of a body under the action of gravity, and say that any system in equilibrium under the influence of any force (mechanical, electrical, magnetic, chemical, &c.) whatever, is in stable or unstable equilibrium according as, when slightly disturbed from the position of equilibrium, its potential energy is increased or decreased by the displacement. If when disturbed the potential energy of the system remains the same, then the equilibrium is neutral.

CHAPTER XIV

THE PENDULUM

112. Simple Pendulum.—A heavy particle suspended by a perfectly flexible weightless thread forms what is called a simple pendulum. Although it is impossible to realise a simple pendulum, we may closely approach the required conditions if we suspend a small metal sphere by a long and very thin thread. The distance between the point of support and the centre of the metal sphere will then be the *length* (l) of the simple pendulum.

Let OA (Fig. 86) be a simple pendulum in its position of rest. Suppose the pendulum deflected from its position of rest to the position OB. If m is the mass of the pendulum bob, then we have a force mg acting through B vertically downwards, *i.e.* parallel to OA. This force may be resolved into a component $mg \cos \theta$ along OB, where θ is the angle AOB, and a component $mg \sin \theta$ along BC, the tangent to the circular arc along which the bob moves. The first of these components, namely that along BD, does not tend to bring the pendulum back to its equilibrium position, but simply causes a tension in the supporting thread. The other component, $mg \sin \theta$ along BC, on the other hand; tends to bring the pendulum back to its undisturbed position. Since mg is constant, whatever the displacement, it follows that the force of restitution is proportional to the sine of the angle through which the pendulum is displaced. The distance through which the pendulum has been displaced is the length of the arc AB, and is equal to $l\theta$, where l is the length (OB) of the pendulum. The ratio of the force of restitution to the displacement is therefore

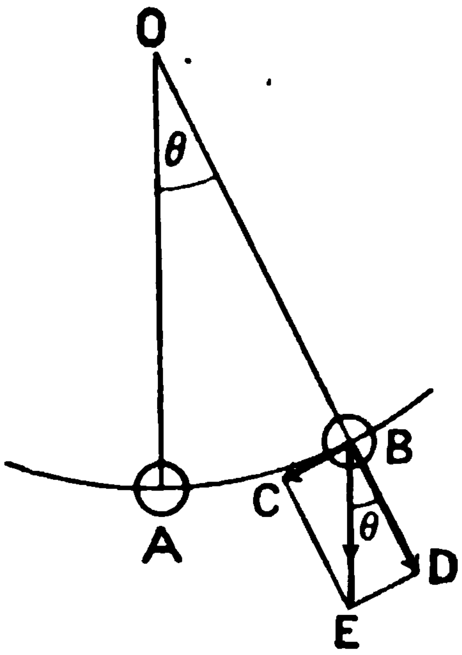


FIG. 86.

$$\frac{mg \sin \theta}{l\theta} = \text{constant} \times \frac{\sin \theta}{\theta}.$$

Now it has been shown in § 51 that when a body moves in a S.H.M. the acceleration is proportional to the displacement. Hence, as the force acting on a body is proportional to the acceleration it produces

(Newton's second law), it follows that the force which must act on a body to cause it to execute a S.H.M must be proportional to the displacement. Thus if a body moves from rest so that the force tending to bring it back to its position of rest is proportional to its displacement, or that the ratio of the force of restitution to the displacement is constant, the body will execute a simple harmonic motion. Hence if the quantity sin θ/θ is constant a pendulum will execute a S.H.M.

The following table gives some values of the ratio sin θ/θ for different values of θ:—

Degrees.	Radians $=\theta$.	Sin θ.	Sin θ/θ.
0° 10′	.002909	.002909	1.0000
0° 30′	.008727	.008727	1.0000
1° 0′	.017453	.017452	1.0000
2° 0′	.034907	.034900	.9998
3° 0′	.052360	.052336	.9995
10° 0′	.174533	.173648	.9949
20° 0′	.349066	.342020	.9798

It will be seen that for values of θ up to 2° or 3° the quotient sin θ/θ is constant to within one part in 5000. The motion of a pendulum is therefore a S.H.M. as long as the angle through which it swings is not too great. The table also shows that the value of sin θ/θ decreases as the angle θ increases; this means that for large displacements the force of restitution increases more slowly than the displacement, and hence the pendulum will take longer to complete a vibration when the displacement is large than it does when the displacement is small.

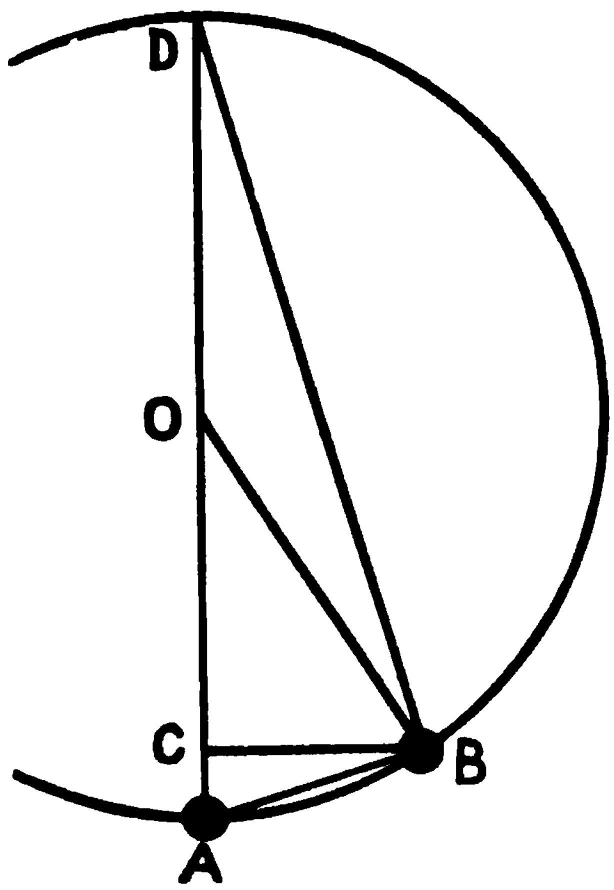


FIG. 87.

113. Time of Oscillation of a Simple Pendulum.—In § 51 it has been shown that in a S.H.M. of amplitude a the maximum velocity is ωa, where ω is the angular velocity in the circle of reference. Since $\omega = 2\pi/T$, where T is the periodic time of the S.H.M., the maximum velocity is equal to $2\pi a/T$. Hence the kinetic energy of the simple pendulum, of which the bob has a mass m when it passes through its position of equilibrium, is $2\pi^2a^2m/T^2$. If we consider that the pendulum has no potential energy when it is in its position of equilibrium, the whole energy when it passes through this point is kinetic, and is equal to $2\pi^2a^2m/T^2$. When the pendulum is at the extremity of its swing it is momentarily at rest, hence

it possesses no kinetic energy, the whole of the energy being potential energy.

Let O (Fig. 87) be the point of suspension of the pendulum, OA the position of rest, and OB the position of maximum positive elongation. Then when the bob is at B it has been raised against gravity through a height $\overline{AC}$, hence its potential energy is $mg\,\overline{AC}$ (§ 76).

Since the triangles ABC, ADB are similar,

$$\frac{\overline{AC}}{\overline{AB}}=\frac{\overline{AB}}{\overline{AD}}$$

$$\therefore \qquad \overline{AC}=\frac{\overline{AB}^2}{\overline{AD}}.$$

Since

$$\overline{AD}=2AO=2l,$$

we have

$$\overline{AC}=\frac{\overline{AB}^2}{2l}$$

Hence the potential energy when the bob is at rest at the extreme elongation is

$$\frac{mg\overline{AB}^2}{2l}$$

The whole of the energy of the vibrating pendulum being kinetic at A and potential at B, and since the energy at A must be equal to the energy at B, we get

$$\frac{mg\overline{AB}^2}{2l}=\frac{2\pi^2a^2m}{T^2}.$$

In this equation a is the arc AB, and $\overline{AB}$ is the chord of this arc. If the amplitude of the vibrations are sufficiently small, the chord may be taken as equal to the arc, and then

$$g/2l=2\pi^2/T^2,$$

or

$$T=2\pi\sqrt{\frac{l}{g}}$$

It must be remembered that this expression only holds, and the vibrations are only isochronous, *i.e.* the periodic time T independent of the amplitude, when the amplitude of the pendulum is small.

Another method of obtaining the expression for the period of a simple pendulum is as follows. It has been shown in § 51 that in the case of a S.H.M., when the displacement is d the acceleration is $d\omega^2$, where ω is the angular velocity in the circle of reference. The force acting, which is

equal to the product of the mass into the acceleration, is therefore equal to $md\omega^2$, where m is the mass of the pendulum bob.

It has been shown in § 112 that the force of restitution, when the displacement is $l\theta$, is $mg \sin \theta$; hence, putting $l\theta$ for d in the expression for the force of restitution in the preceding paragraph, and then equating the two expressions, we get

$$ml\theta\omega^2 m = mg \sin \theta ;$$

or, since $\omega = 2\pi/T$, where T is the period of the vibration,

$$\frac{4\pi^2 l\theta}{T^2} = g \sin \theta \cdot$$

$$\therefore \qquad T = 2\pi \sqrt{\frac{l}{g} \cdot \frac{\theta}{\sin \theta}} \cdot$$

If the amplitude of the vibration is so small that $\theta / \sin \theta = 1$, we get

$$T = 2\pi \sqrt{\frac{l}{g}},$$

the expression found before. The numbers given in the table on p. 129 show that, if the amplitude does not exceed 3°, the above expression gives the time of vibration correct to about 1 in 5000.

114. The Compound Pendulum.—If it were possible to construct a pendulum with a sufficiently small bob, and a suspending thread so thin and flexible that it could be regarded as a simple pendulum, we might employ it to measure g, the acceleration of gravity. For, rewriting the expression obtained in the last section, we get

$$g = \frac{4\pi^2 l}{T^2}.$$

Hence, if the length (l) and the periodic time (T) of a simple pendulum are observed, we may calculate g.

Although it is physically impossible to realise a simple pendulum, it is possible to determine what would be the length of the simple pendulum, which would vibrate with the same period as a pendulum of any given form; this operation is called finding the length of the equivalent simple pendulum.

We may look upon any actual pendulum as built up of a number of material particles. Each of these particles, since it is at a fixed distance from the point of support, tends to oscillate as a simple pendulum with its own proper period; but as the distances between the various particles and the point of support are different, the periods in which the particles tend to oscillate are different. However, since they are all rigidly con-

nected together, they are all obliged to oscillate with the same period. Thus some of the particles will be obliged to oscillate faster than their natural period, while others will be obliged to oscillate more slowly. There will, however, always be at least one particle which will oscillate in its own natural period, and the distance between this particle and the point of support will be the length of the equivalent simple pendulum. That point in a pendulum where the particles are oscillating with the same period as they would have if they were the bobs of simple pendulums suspended from the same point of support as the actual pendulum, is called the *centre of oscillation*, the point of support being called the *centre of suspension.*

From what has been said, it will be seen that if the whole mass of the pendulum is supposed to be concentrated at the centre of oscillation, so as to form a simple pendulum, then the period of this simple pendulum would be the same as that of the compound pendulum. For many purposes it is convenient to consider that the whole mass of a compound pendulum is concentrated at its centre of oscillation.

There are many very interesting properties of the centre of oscillation —to prove which would involve more mathematics than it is possible to introduce into these pages—which may with advantage be stated here: (1) The centre of oscillation and the centre of suspension are interchangeable. Thus the period of a compound pendulum will be unaltered if it is suspended from its centre of oscillation, and the old centre of suspension will then be the new centre of oscillation. (2) If the pendulum is struck a blow at its centre of oscillation it will rotate round the centre of suspension, but the blow will not produce any pressure on the axle or knife-edge by means of which the pendulum is supported. For this reason the centre of oscillation is often called the *centre of percussion.* A familiar example of the above property of the centre of oscillation occurs when a ball is struck with a cricket-bat. If the ball is struck with a certain part of the bat, the impact is not felt by the hands of the striker; if, however, the bat is struck either higher up or lower down, a distinct "sting" is felt in the hands. When no sting is felt, the ball has been struck with the part of the bat which is the centre of oscillation corresponding to the centre of supension where the bat is held in the hands. (3) If a body is so placed that it is free to turn in any way, say by being floated on water, and is struck at a point A; then if, under the influence of the blow, it moves so that a point B remains at rest, the points A and B will be to one another in the position of centre of suspension and centre of oscillation.

115. Kater's Reversible Pendulum.—Captain Kater in 1818 made use of the fact that the centres of suspension and oscillation are interchangeable, to obtain the distance between these centres, or, in other words, the length of the equivalent simple pendulum, and from this the value of the acceleration of gravity (g). Kater's pendulum consists of a

metal rod AB (Fig. 88), which carries at one extremity a heavy lens-shaped mass C. Two steel knife-edges N_1 and N_2 are fixed to this rod with their edges turned towards one another, and at such a distance that the pendulum vibrates with very nearly the same period whether it is suspended from one or the other. In order to allow of the time of vibration about these two knife-edges being exactly adjusted to equality, two movable metal pieces (D and E) fit round the rod. The heavier of these pieces (D) can be firmly fixed to the rod by means of clamping screws, and the lighter piece (E) is attached to the other by a fine screw. The two pieces are moved till the adjustment to equal periods is very nearly complete, then the position of the lighter piece (E) is finally adjusted by means of the screw till the periods are exactly equal. When this adjustment is complete, the distance between the knife-edges is equal to the length of the equivalent simple pendulum, which would vibrate with the same period as the reversible pendulum. This time of vibration is then very carefully determined by comparison with a clock, the rate of which is itself determined by astronomical observations on the stars made with a transit instrument.

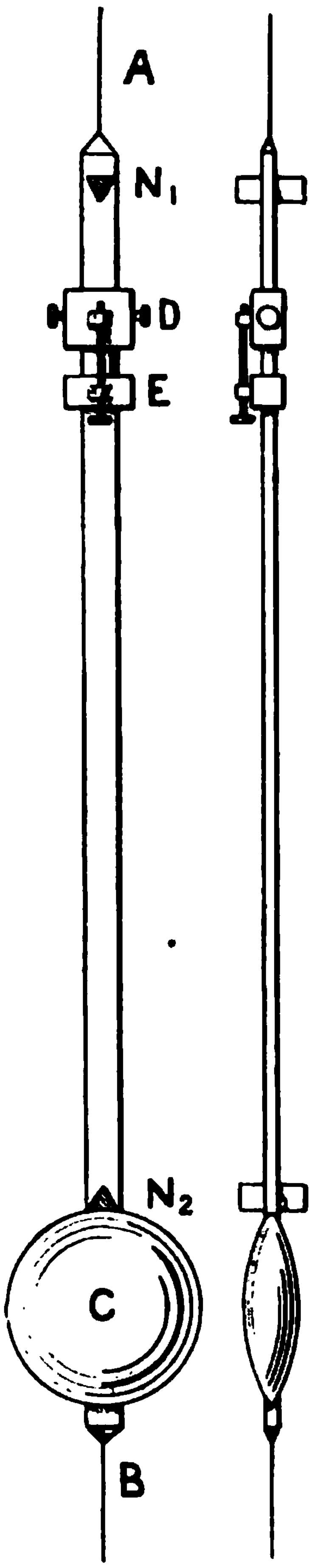


FIG. 88.

116. Variations in the Value of g at Different Parts on the Earth's Surface.—If the earth were a perfect sphere of uniform density, or, at any rate, if it were built up of spherical shells, each shell being of uniform density, the attraction exerted by the earth on a given mass would be the same at all points on the earth's surface. Since, however, the earth is really a spheroid, the diameter through the poles being about 42 kilometres less than the minimum equatorial diameter, the value of g is greater at the poles than at the equator. The fact that the earth, and therefore any body on the surface of the earth, is in rotation about the polar axis, also makes the value of g less at the equator than at the poles; the reason being that, as a body on the earth's surface is moving in a circle, part of the force exerted by the earth's gravitational attraction is required to keep the body from flying off along a tangent, *i.e.* is used in neutralising the so-called centrifugal force.

The amount by which the weight of a body at the equator is less than the force with which the earth attracts it, can be easily calculated, as follows:—

If R is the radius of the earth, then the force along the radius neces-

sary to keep a body having a mass of 1 gram moving in a circle of radius R is (see § 42) $\left(\frac{2\pi R}{T}\right)^2/R = \frac{4\pi^2 R}{T^2}$ dynes.

Now $R = 6.37 \times 10^8$ cm., and $T = 86164$ sec. (one *sidereal* day).

Hence the "centrifugal force" $= \frac{4\pi^2 6.37 \times 10^8}{(86164)^2} = 3.4$ dynes.

Since at the equator the value of g is 978.0 cm./sec.^{2}, or the weight of body having a mass of 1 gram is 978.0 dynes, it follows that if the earth were at rest the weight of the body would be 981.4 dynes.

At the poles the weight of a body is unaffected by the rotation of the earth, and hence g is not diminished on this account.

According to Helmert, the value of g at a place in latitude ϕ and at the sea-level is given by the equation

$$g = 977.989(1 + 0.0052 \sin^2\phi) \text{ cm./sec.}^2,$$
$$\text{or} = 32.0862(1 + 0.0052 \sin^2\phi) \text{ foot/sec.}^2.$$

The following table gives the value of g at a few places :—

	Latitude.	Value of g in	
		Cm./sec.^{2}.	Foot/sec.^{2}.
Equator	0	977.99	32.086
Latitude 45° . . .	45°	980.60	32.172
Paris	48°50′	980.96	32.184
Greenwich . . .	51°29′	981.17	32.191
Berlin	52°30′	981.26	32.194
Dublin	53°21′	981.32	32.196
Manchester . . .	53°29′	981.34	32.196
Edinburgh . . .	55°57′	981.54	32.203
Aberdeen . . .	57°9′	981.64	32.206
Pole	90°0′	983.21	32.258

117. Gravity Independent of the Nature of the Matter.—The quantity g in the expression found in § 113 for the time of vibration of a simple pendulum, expresses the acceleration with which the bob would fall *in vacuo* under the influence of the earth's attraction. For, in obtaining that formula, we called the force with which the earth attracts the bob mg; in other words, we have measured the force by the product of the mass of the body into the acceleration it produces. Suppose, however, that equal masses of two different substances, say brass and stone, as defined according to Newton's second law (see § 60), are not equally attracted by the earth, so that the gravitational force between two bodies depends not only on their mass, but on the nature of the substance of which the bodies are formed. Under these circum-

stances, the value of g obtained by pendulum observations from the equation

$$g=4\pi^2 l/T^2$$

would depend on the nature of the material of which the pendulum bob is formed. That such a difference might quite well be expected, will be seen if we remember that equal masses are defined to be such that a given force produces in them equal accelerations, while in the case of many other cases of attraction exerted between bodies, such as the force exerted on a body by a magnet, the magnitude of the force depends on the nature of the material of which the body is composed.

The first to investigate whether the *weight* of a body was proportional to its mass was Newton. He employed in the first place two pendulums of the same length (l), and having bobs made of the same material but of different mass (m_1 and m_2). Suppose g_1 and g_2 are the accelerations which the earth's attraction would produce in the masses m_1 and m_2 respectively, then the force (F_1) with which the earth attracts the first mass is $m_1 g_1$, and that (F_2) with which it attracts the second mass is $m_2 g_2$. Hence

$$g_1=\frac{F_1}{m_1} \text{ and } g_2=\frac{F_2}{m_2}.$$

But by the law of the simple pendulum

$$g_1=\frac{4\pi^2 l}{t_1^2} \text{ and } g_2=\frac{4\pi^2 l}{t_2^2},$$

where t_1 and t_2 are the periodic times of the pendulums. Equating the expressions for g_1 and for g_2, we get

$$\frac{F_1}{m_1}=\frac{4\pi^2 l}{t_1^2} \text{ and } \frac{F_2}{m_2}=\frac{4\pi^2 l}{t_2^2}.$$

Hence if, as Newton found was the case, the periodic times of the pendulums are the same, *i.e.* if $t_1=t_2$, then

$$\frac{F_1}{m_1}=\frac{F_2}{m_2} \text{ or } \frac{F_1}{F_2}=\frac{m_1}{m_2},$$

and therefore the forces exerted by the earth on two bodies of the same material (*i.e.* the weights) are to one another as the masses of the bodies.

Newton then proceeded to experiment with pendulums the bobs of which were made of different materials, but which were all of the same length, and he was unable to detect any difference in the periodic times, so that the value of g was the same whatever the material of the bob. These experiments were afterwards repeated in a much more accurate manner by Bessel, in 1832. Bessel used, as the bob of his pendulum, a hollow brass cylinder which could be filled with the different materials to be tested. By this means the effect of the resistance of the air on the

time of oscillation of the pendulum remained the same, since the surface of the bob was the same in all the experiments.

Bessel found that the values of g, as determined with pendulums of brass, but of different masses, or with pendulums of various materials, such as iron, zinc, lead, silver, gold, meteoric iron, marble, quartz, water, &c., did not differ one from another by more than about 1 part in 70,000. He further showed, by making repeated observations with the same pendulum, that the differences obtained were of about this magnitude, so that such small variations in the value of g as were obtained must be set down to experimental error.

118. The Pendulum as a Measure of Time.—The property of the pendulum that, so long as the maximum amplitude is small, the period of vibration is independent of the amplitude of the vibrations, has received a most important practical application in the employment of the pendulum in the measurement of time. In order to employ a pendulum as a timekeeper, we require to supplement it by some mechanism which shall, in the first place, count the number of oscillations of the pendulum; and, in the second place, keep the pendulum swinging by supplying enough energy to allow for the energy which it loses in friction at its point of support and against the air. The necessary energy is stored up in a raised weight, a bent spring, or an electric battery, and is doled out to the pendulum at the proper rate by the mechanism which is employed to count the oscillations, and which is called the escapement. Although space will not permit of describing the forms of escapement in general use, we may, by a consideration of the laws of the pendulum, arrive at the characteristic which should distinguish a good escapement.

Suppose that a pendulum OA (Fig. 89) is swinging so that B and D are the positions of maximum elongation, and that at the moment when the pendulum is passing through its position of rest (A), moving from right to left, we give the bob a sudden blow in the direction of the arrow. The result will be that the pendulum will start its elongation on the left side with a greater velocity than it would otherwise do, but it will reach its new position of *maximum* elongation C in exactly the same time as it would have reached its old position of maximum elongation B, since we have only increased the amplitude of the vibrations, and the time of vibration is independent of the amplitude.[1] Next, suppose that the blow is delivered when the pendulum is at its position of maximum elongation D. It will now move from D with the velocity it would have had if it had come from some point such as E, and will therefore

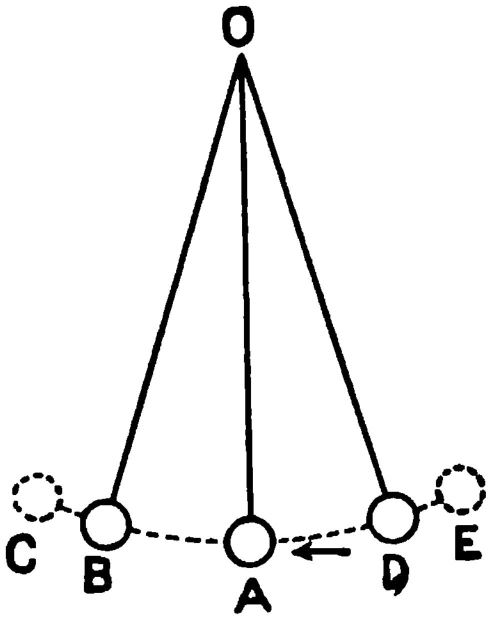


FIG. 89.

[1] The amplitude being of course small, and nothing like as large as that shown in the figure.

reach its position of rest A in less than a quarter period, for it has moved from its position of maximum elongation D to its position of rest A with a velocity greater than it would naturally possess suppose no blow had been struck. The effect, therefore, of striking the pendulum a blow when at D is to accelerate its period of vibration for the next quarter period. In the same way, if the blow had been struck in the opposite direction it would have retarded the motion of the pendulum. Thus it can be shown that if we attempt to keep up the movement of the pendulum by supplying it with energy, we must only do this when it is passing through its position of rest, for if we interfere with the free swing of the pendulum during any other part of its swing the period of the pendulum will be affected. Hence a good escapement is so arranged as to give a small impulse to the pendulum each time it passes through its position of rest, but to allow the pendulum to move quite freely during the rest of its swing.

119*. Bifilar Pendulum.—If a rigid body, such as a metal bar, AB (Fig. 90), is suspended by means of two strings CD and EF of equal length l, so that the centre of gravity G lies between A and B; then, so long as there is no torsion in the suspending strings, the bar will take up a position of equilibrium such that CD and EF are both in the same vertical plane. Such an arrangement is called a bifilar pendulum, and when the bar AB is twisted about a vertical axis, it will be acted upon by a couple tending to restore it to its position of rest. In order to find the magnitude of the restoring couple, we will suppose that $\overline{CE}$ is equal to $\overline{DF}$, *i.e.* that when the bar is in its position of equilibrium the suspending strings are parallel. Further, we shall suppose that the centre of gravity of the bar lies exactly half-way between D and F, so that half the weight of the bar is carried by each string.

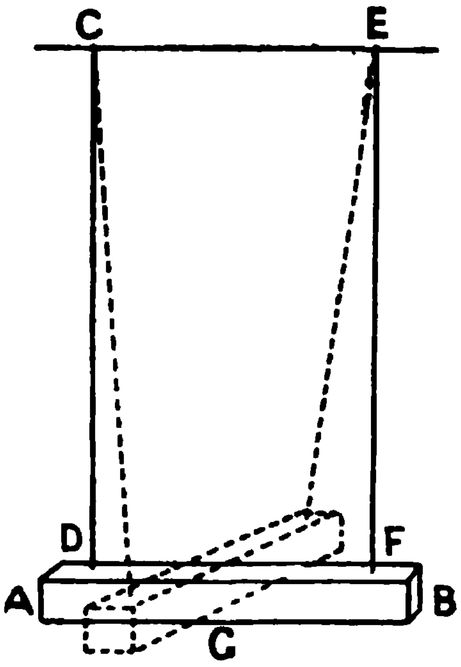


FIG. 90.

If then the mass of the bar is m, we may, if we like, consider that we have two masses, each equal to $m/2$, one at D and the other at F, and that they are rigidly connected by a weightless rod $\overline{DF}$. Let this rod $\overline{DF}$ (Fig. 91) be twisted through an angle θ, so that each of the suspending threads now makes an angle ϕ with its former position. Considering one end only, let CD (Fig. 92) be the undisturbed position of one of the suspending threads, and CD′ the disturbed position. Then we have a force of $\frac{m}{2}g$ acting at D′ vertically downwards. This force must be resolved horizontally and along CD′ produced. In the triangle D′KH the angle D′HK is equal to ϕ. Hence the component along D′K is equal to $\frac{m}{2}g\frac{D'K}{D'H}=\frac{m}{2}g\tan\phi$.

This force acts along D_1D_1' in the right hand part of the figure, which represents a plan of the arrangement. Therefore the moment of the force about G is

$$\frac{m}{2}.g \tan \phi . \overline{GK}.$$

But
$$\overline{GK} = \overline{D_1G} \cos \frac{\theta}{2} = \frac{a}{2} \cos \frac{\theta}{2},$$

where a is the distance DF between the suspending threads.

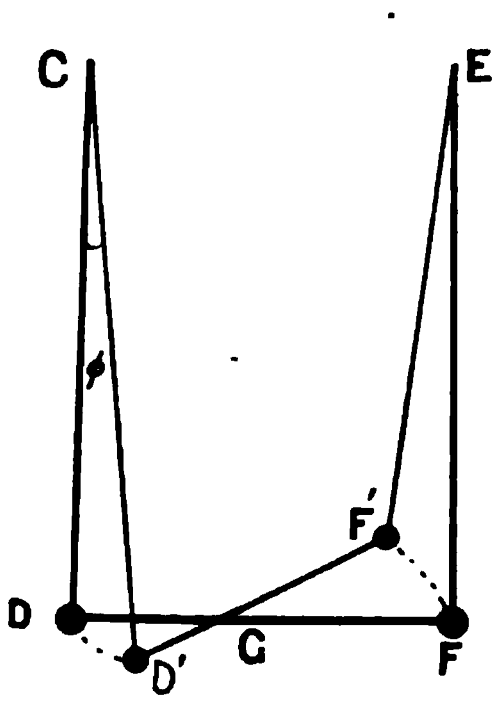


FIG. 91.

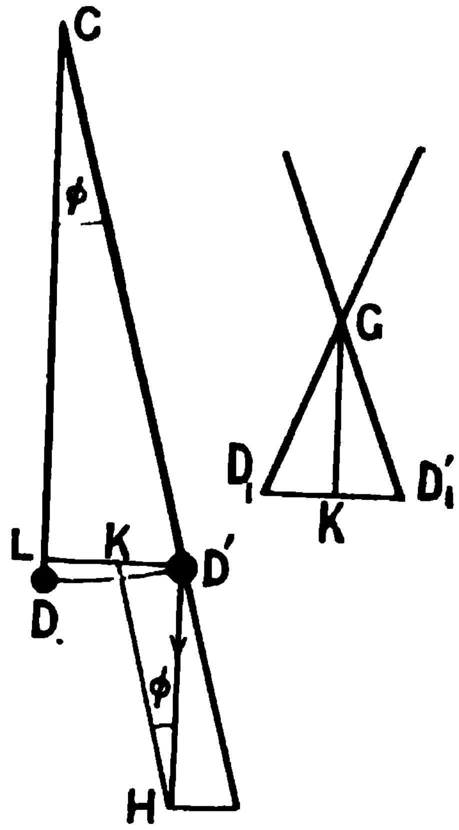


FIG. 92.

Hence the moment about G is

$$\frac{mga}{4} \cos \frac{\theta}{2} . \tan \phi.$$

Now $\tan \phi = \frac{\overline{LD'}}{\overline{CL}} = \frac{\overline{D_1D_1'}}{\overline{CL}} = \frac{2\overline{D_1K}}{\overline{CL}} = \frac{2a \sin \frac{\theta}{2}}{\overline{CL}}$, and for small deflections, *i.e.* small values of ϕ, $\overline{CL}$ is very nearly equal to CD, *i.e.* l.

Hence the restoring moment is equal to

$$\frac{mga^2}{2l} \sin \frac{\theta}{2} \cos \frac{\theta}{2},$$

or
$$\frac{mga^2}{4l} \sin \theta \text{ ; since } 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} = \sin \theta.$$

Therefore the total restoring moment, taking into account the two ends, is

$$\frac{mga^2}{2l} \sin \theta.$$

From this expression we see that the moment of the restoring forces is

proportional to the sine of the angle of deflection (θ); hence, just as in the case of the simple pendulum, so long as θ is small, and therefore $\sin\theta$ practically proportional to θ, the system will execute oscillations which form a simple harmonic motion.

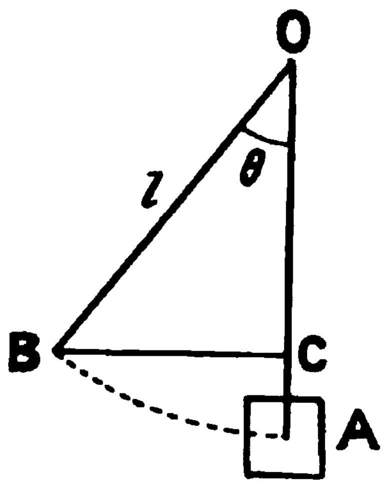


FIG. 93.

The bifilar pendulum can also be used to measure experimentally the moment of a couple, for if the couple produce a twist of θ, then the moment is given by the expression found above.

120. Ballistic Pendulum.—Suppose we have a heavy block A (Fig. 93) of mass M suspended by a thread of length l from a point O, and that a bullet of mass m moving in a horizontal direction with a speed v strikes A opposite its centre of oscillation, and sinks into the block. From what has been said in § 114, it follows that under these circumstances the block A will move in a circle with O as centre, and will not be set into rotation about a vertical axis. Let B be the maximum elongation of the block, then the work done against gravity while the block, with the bullet imbedded, has moved from its position of rest to its maximum elongation is $(M+m)g\,.\,\overline{CA}$. This work has been done at the expense of the kinetic energy of the bullet; hence

$$\tfrac{1}{2}mv^2 = (M+m)g\,.\,\overline{CA}.$$

But

$$\overline{CA} = \overline{OA} - \overline{OC}$$
$$= l - l\cos\theta = l(1-\cos\theta).$$

Hence

$$\tfrac{1}{2}mv^2 = (M+m)gl(1-\cos\theta),$$

or

$$v = \sqrt{\frac{2gl(M+m)(1-\cos\theta)}{m}},$$

so that, by observing the angle θ, the velocity of the bullet when it struck the block can be calculated.

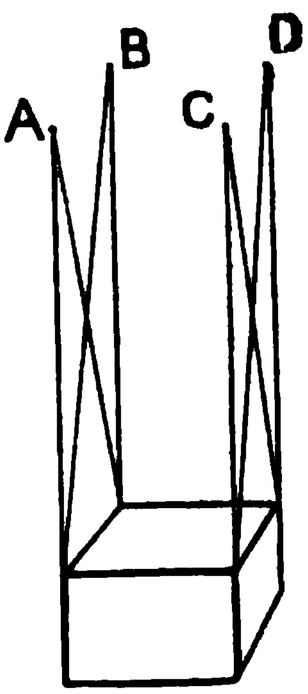


FIG. 94.

In order to prevent the possibility of the block being set in rotation round a vertical axis, and hence the utilisation of part of the kinetic energy of the bullet in setting up this rotation, it is usual to suspend it by eight strings, as shown in Fig. 94. When thus suspended the block can only turn about an axis parallel to $\overline{AB}$ or $\overline{CD}$.

PART IV.—PROPERTIES OF MATTER

CHAPTER XV

PROPERTIES OF MATTER

121. General Properties of Matter.—We have, in § 2, defined matter as that which can occupy space, and in the subsequent pages we have dealt with certain of the more general properties which matter possesses. Thus we find that matter always exists in a definite quantity, so that we say that, as far as we are able to tell, every body has a definite mass. Another property of matter which is quite general, being shared by all kinds of matter alike, is *extension*, *i.e.* every piece of matter must occupy some definite volume of space, and in fact we have used this property to define matter.

Another property of matter intimately related to the last is that of impenetrability, or that two portions of matter cannot occupy the same space. Since, however, under certain circumstances, the volume of a mixture of two kinds of matter is sometimes less and sometimes greater than the sum of the volumes of the two constituents, we are obliged to introduce certain restrictions into the statement that matter is impenetrable.

If 50 c.c. of alcohol is mixed with 50 c.c. of water, the volumes of the mixture, instead of being 100 c.c., is only about 97 c.c. It would thus appear that the water and alcohol are able to a certain extent to interpenetrate.

We may, however, regard matter as built up of a number of small masses called molecules, and suppose that even in the case of a solid or liquid these molecules are not in actual contact, but that there exist interspaces which are unfilled by matter. Thus, although the molecules may be rigorously impenetrable, yet under certain circumstances we may suppose, either that fresh molecules can be packed in the interstices, or that the molecules can be packed closer together, and hence the volume of the body as a whole may be reduced.

That bodies in general are porous, *i.e.* that there exist pores or channels in bodies, which to the eye, and even in some cases under the highest powered microscope, appear quite solid and impervious, is evident by many experiments. Thus Bacon, when he attempted to compress water by squeezing a lead sphere filled with water, found that the water exuded through the pores of the lead and stood in beads on the surface. In the somewhat more famous experiment made by the Florentine

Academicians, in which the water was enclosed in the first place in a silver sphere, and thereafter in a silver sphere which had been thickly gilded, the same result was obtained, the water passing through the silver and gold. In a later section, when dealing with osmotic pressure, we shall see what use may be made of the porosity of certain materials.

Another property of matter is divisibility, that is, all such portions of matter as we are acquainted with can be subdivided. The interest of the study of this property of matter arises when we consider whether it would be possible, supposing we had suitable means, to continue subdividing a given piece of matter, say gold, indefinitely, or whether we should finally arrive at such a condition that if we subdivided the excessively minute piece of gold under test, the portions would no longer possess the properties of gold, but be some new kind of matter of which we may consider gold to be built up. Considerations of the relations between the quantities of different kinds of matter which are found by chemical analyses to build up all known compounds, seem to point conclusively to the conclusion that matter is *not* infinitely divisible, but that we should finally arrive at the ultimate particle of the given kind of matter, so that if we did succeed in carrying the subdivision any further, we should no longer have this kind of matter, but should have split it up into its components. Hence we suppose that for every kind of matter there exists some particle which is the smallest that can possibly exist.

Although we cannot hope to prepare such an ultimate particle by physical subdivision, it is interesting to note some of the examples of extreme tenuity which have been obtained. Ordinary gold-leaf, which is prepared by beating out a small nodule of gold between two pieces of skin, has a thickness of about 1/300,000 inch, or 8.2×10^{-6} centimetre; and Faraday has shown that the thickness can be very much further reduced by immersing it in cyanide of potassium, which dissolves the surface. It has been found possible to rule parallel lines on a polished piece of glass at a distance of 1/100,000 inch (10^{-5}) or 2.5×10^{-5} centimetre from one another. By melting a small piece of quartz in an oxy-hydrogen blowpipe, it is possible to prepare threads of quartz having a diameter of less than a hundred thousandth (10^{-5}) of an inch, or 2.5×10^{-5} centimetre.

The whole science of chemistry, as far at any rate as it is a quantitative science, is based on the supposition that by no possible means can the quantity of matter be altered, so that if the mass of the matter within a given space is found to alter it must be due to matter having passed into or out from the given space through its boundary. Chemistry also teaches that, as far as the elements are concerned, the quantity of *each* element in any given space cannot be altered except by the introduction or withdrawal of such element through the boundary of the space. Hence matter is said to be indestructible, *i.e.* we can neither create it nor destroy it, nor can we convert one element into another.

The fact that Newton's first law of motion is found to apply to all kinds of matter shows that inertia is one of the general properties of matter. The universal truth of Newton's law of gravitation, or the fact that at the surface of the earth all matter possesses weight, constitutes a property of matter.

In addition to the above properties, which are common to matter in all its forms, and are therefore called general properties of matter, there are other properties which may be quite different for the different kinds of matter, or for the same kind of matter when its surroundings or state are different, and are consequently called contingent properties of matter. We shall consider these properties in the appropriate sections under the headings of sound, light, &c., where the groups of phenomena depending on each particular property are dealt with.

122. Elasticity.—When a system of forces acts on a portion of matter, although they may not produce any motion of the body as a whole, yet they may produce a displacement of the various particles of the body relatively one to the others. Such a relative motion of the different portions of a body is called a *strain*. In the preceding pages, when dealing with the action of forces on a body, we have supposed the body to be rigid; that is, we have supposed that, however great the forces acting on the body, they did not produce any strain. The subject of the strains produced in different kinds of matter under various conditions can, in such a work as the present, be only dealt with in a very rudimentary manner.

Strains may be divided into two classes, according as to whether they consist of a change in volume of the body or a change in shape.

When a body is strained by the application of external forces it in general resists the strain, forces being called into play by the relative displacement of its parts tending to cause it to return to its unstrained condition. This restoring force, called into play owing to the strain, is called a *stress*.

It is important to note that the magnitude of the stress induced in a body is not necessarily always equal to that of the forces applied to produce the strain. Thus it requires a considerable force to deform or strain such a body as a lump of moist clay; but if the external forces are removed, there is no tendency for the clay to regain its former shape. If, however, the deforming forces strain the body till the stress induced is such as just to prevent further strain, then a state of equilibrium will be set up, and the stress will be exactly equal and opposite to the deforming force.[1]

A body which offers a stress, tending to restore the body to its original condition when it is strained, is said to be elastic or to possess elasticity.

[1] As we shall only deal with the cases where a strained body is in a state of equilibrium, the deforming forces will always be equal and opposite to the stress called into play. Hence it saves circumlocution if we use the term stress to indicate the deforming forces.

A body in which change of volume calls into play a stress is said to have *volume elasticity*.

A body in which change of shape, without change of volume, calls into play a stress is said to possess *rigidity*.

The elasticity of a body is measured by the ratio of the stress produced by a given strain to that strain, or

$$\text{Elasticity} = \frac{\text{stress}}{\text{strain}}.$$

The question as to how the strain and stress have to be measured in different cases is postponed till we come to consider in greater detail the properties of the various forms of matter.

123. States of Matter.—For the purposes of subdivision we may say that matter exists in three distinct states, the solid, the liquid, and the gaseous. In addition, however, to states which fulfil the definitions of a solid, a liquid, or a gas, which we shall give later on, it will be found that there are intermediate states which bridge over the intervals between the solid and the liquid, and the liquid and the gas. As an example of the kind of gradation which exists, we may take the following: steel, lead, wax, cobbler's-wax (which will flow like a liquid if allowed sufficient time), treacle, water, ether, liquefied carbon dioxide, steam, sulphur dioxide, air, hydrogen. In addition there is the critical state when a substance is to all intents and purposes both a liquid and a gas, and the state of extreme rarefication of a gas which is sometimes called the radiant state or "fourth state" of matter.

We may define a solid as a portion of matter which is able to support a steady longitudinal stress without lateral support. In contradistinction, a portion of matter which is unable to support a steady longitudinal stress without lateral support is called a *fluid*.

If we take a solid body, say a lead-pencil, then we may apply a deforming force, either of compression or extension, in any direction to the pencil, and there will be a certain amount of strain, either elongation, compression, or bending produced, which will call into play a stress that, so long as the deforming force is not too great, will be in equilibrium with this force, and this stress will be produced without the body being supported in any way in a direction at right angles to that along which the stress acts. In the case of a fluid, such as water or air, we are unable to exert a stress on it, and hence produce a corresponding strain, unless we supply some constraining boundary which shall prevent the fluid swelling out at right angles to the line of action of the stress. Thus if we have a fluid enclosed in a cylindrical tube between two pistons, then we may apply a deforming force to the fluid either by forcing the pistons towards one another, or by pulling them apart, in one case producing a compression, and in the other a tension in the direction of the axis of the tube, and a stress will in both cases be produced in an opposite direction to

the applied force. If, however, part of the wall of the tube between the pistons is removed, and we then attempt to apply stress to the liquid, we shall not succeed, for either the liquid will flow out through the gap in the tube, or air will be sucked into the tube through this opening, and the fluid will remain unstrained. It is only, therefore, when the column of fluid is laterally supported by the walls of the tube that it is capable of exerting a longitudinal stress.

Fluids are divided into liquids and gases. A liquid is a fluid such that when a certain volume is introduced into a vessel of greater volume it only occupies a portion of the vessel equal to its own volume. A gas is a fluid such that if a certain volume is introduced into a vessel, then, whatever the volume of the vessel may be, the gas will distribute itself throughout the vessel.

124. The Constitution of Matter.—The question as to the finite divisibility of matter has been referred to in § 121. The theory that matter is not infinitely divisible, but that every body is made up of small indivisible parts called atoms (from ἄτομος) is one of extreme antiquity. At the present day this theory is generally accepted, and we are hence led to consider what is the nature of these atoms. In chemistry it is usual to apply the term molecule to the smallest portion of any kind of matter which can exist alone and still preserve the character of that kind of matter, and to restrict the term atom to the smallest portion of any element which can take part in a chemical reaction. In the case of such a substance as chalk, the molecule is the smallest portion of chalk which can exist.

The *molecule* of *chalk* can, by certain processes, be further subdivided, but the parts have no longer any of the attributes of chalk ; they may be carbon dioxide, and lime. These again can be split up into carbon, oxygen, and calcium, but further than this it has up to now been impossible to go. For this reason chalk or carbonate of lime, carbon dioxide, and lime are called compounds, since the molecule of these bodies can be further subdivided, losing, however, in the process their essential properties as chalk, carbon dioxide, &c. On the other hand carbon, oxygen, and calcium are called elements, since the molecule of these bodies cannot by any known means be split up into any simpler bodies. For the purposes of the physicist, as distinct from the chemist, it is generally unnecessary to distinguish between the smallest particle which can exist of a compound or of an element. For our purposes, in considering the structure of matter, we shall call the ultimate particle a molecule, and shall not, in most cases, further consider the question whether it might not be split up into more elementary molecules.

The original conception of a molecule was that it consisted of a hard sphere, and that bodies were built up of such spheres, which were not necessarily in contact with one another. This conception was further extended by Boscovich, who did away with the consideration that the molecule is a material body occupying a certain space. He considered

the molecule to be a mere mathematical point towards or from which certain forces act. He further supposed that any two molecules attract each other with a force which for considerable distances varies inversely as the square of the distance, but which for small distances becomes changed into a repulsion, which increases as the molecules come nearer and nearer together. The chief difficulty in this theory is that it does not seem capable of explaining the inertia of matter.

One of the most recent theories, and one which very powerfully appeals to the imagination, is Lord Kelvin's vortex atom theory. By vortex motion is meant a form of motion such as occurs in a smoke-ring. The path of the particles of air in such a smoke-ring is indicated by the arrows in Fig. 95, where the curved arrows show the direction in which the air particles, which are simply rendered visible by the smoke, rotate, while the straight arrow shows the direction in which the ring, as a whole, moves. There is a very important difference between this form of motion and a wave motion. In the latter, although the waves travel onwards, the individual particles of the medium in which the wave is being propagated only move through a comparatively small distance from their original position, the motion being handed on from one particle to the next. In vortex motion, however, the particles of the medium themselves move forward, so that in a smoke-ring the particles of air originally forming the ring move on with the ring.

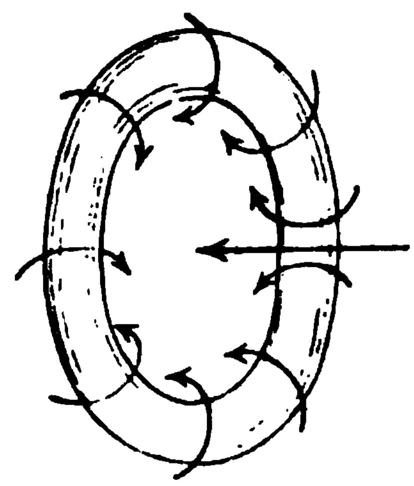


FIG. 95.

The properties of vortex motion were first examined by rigid mathematical methods by von Helmholtz, who found that if the fluid in which this form of motion exists is frictionless, incompressible, and homogeneous, then: (1) A vortex can never be produced, nor if one exists can it be destroyed, so that the number of vortices existing is fixed. This corresponds to the indestructible property of matter. (2) The rotating portions of the fluid forming the vortex maintain their identity, and are permanently differentiated from the non-rotating portions of the fluid. (3) These vortex motions must consist of an endless filament in which the fluid is everywhere rotating at right angles to the axis of the filament, unless the filament stretches to the bounding surface of the fluid. (4) A vortex behaves as a perfectly elastic substance. (5) Two vortices cannot intersect each other, neither can a vortex intersect itself.

On the basis of these results of von Helmholtz, Lord Kelvin has founded a theory as to the constitution of matter. He supposes that all space is filled with a frictionless, incompressible, and homogeneous fluid (the ether), and that an atom is simply a vortex in this medium. The existence of different kinds of atoms may be accounted for by the fact that a vortex need not necessarily be a simple ring, as shown in Fig. 95,

but might have such a form as that shown in Fig. 96. Since a vortex can never intersect itself, it follows that the number of times such a vortex is linked with itself must always remain the same. Hence we may suppose that the atoms of the different elements are distinguished from one another by the number of times they are linked.

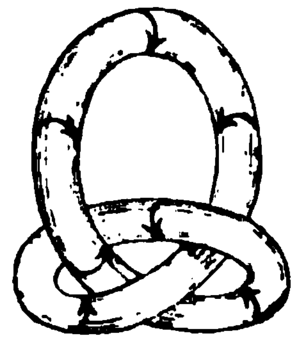


FIG. 96.

125. The Size of Molecules.—Until more is known of the nature of molecules, no very definite statement as to what is meant by the size of a molecule is possible. Since, however, the methods of deducing the size of the molecules at present known only give at the most a rough estimate of the "magnitude" of this quantity, the difficulty of defining what is meant by the size is not very important. For the present it is usual to consider that a molecule consists of a solid sphere, though of course these spheres need not fill even a small fraction of the total space which the body apparently occupies.

The methods by which the following estimates of molecular magnitude have been made cannot be described till the physical phenomena from which they are deduced have been considered.

In the following table the diameters of the molecules of some gases, supposed to be spherical, are given, as well as the mass of a single molecule. Knowing the mass of a molecule and the density, that is, the mass of unit volume at standard temperature and pressure, we may calculate the number of molecules contained in unit volume. Thus if m is the mass of each molecule, and there are n molecules in unit volume, the mass of unit volume, that is, the density d, is given by

$$d = nm.$$

Substituting the values of m from the first column of the table, and the values of the density as given in the table on page 150, it will be found that in each case the value found for n is about 2×10^{19}.

DIMENSIONS OF THE MOLECULES OF GASES.

Gas.	Mass of Molecule.	Diameter of Molecule.
Hydrogen . . .	46×10^{-25} gram.	5.8×10^{-8} centimetres.
Oxygen	736×10^{-25} ,,	7.6×10^{-8} ,,
Carbon monoxide . .	644×10^{-25} ,,	8.3×10^{-8} ,,
Carbon dioxide . .	1012×10^{-25} ,,	9.3×10^{-8} ,,

As a help to the realisation of what the above numbers mean, we may say that seventeen millions of hydrogen molecules, if placed in a row so that one touched the next, would form a row about one centimetre in length. Another illustration has been given by Lord Kelvin, namely: if a drop of water were magnified till it was equal in size to the earth, the molecules would be about the size of cricket-balls.

CHAPTER XVI

PROPERTIES OF GASES

WE commence our study of the general characteristics of the different states of matter with that of the gaseous state, for in this condition we are able to account for most of the observed facts by dynamical reasoning, based on what is known as the kinetic theory of gases. On the other hand, in the case of solids and liquids we are very far from possessing even an approximate dynamical theory to account for the observed properties. The structure of a gas being so much more simple than that of a liquid or solid, it is best to begin by the study of the gaseous state, and then to proceed to that of the liquid and solid states.

Before, however, commencing the study of the special properties of gases it will be convenient to consider some of the general properties of fluids, since these properties are common to both gases and liquids.

126. Pressure Exerted by a Fluid.—Since a fluid cannot resist a stress unless it is supported on all sides, or in other words it has only elasticity of volume, it can offer no resistance to forces which tend only to change its shape and not its bulk.

It follows, from this mobility of fluids, that in the case of a fluid at rest the force it exerts on any surface in contact with it must be perpendicular to the surface. If the force did not act perpendicular to the surface, then it could be resolved into two components, one acting perpendicular to the surface, and the other acting parallel to the surface. This latter component would, if it existed, cause the fluid to move parallel to the surface. Since by supposition the fluid is at rest, and therefore no such tangential motion exists, there can be no tangential component of the force, so that the force exerted by the fluid on the surface is perpendicular to the surface.

The magnitude of a force exerted by a fluid is measured by the force exerted on the unit of surface, and is called the *pressure.*

Hence in the *c.g.s.* system the unit of pressure is a dyne per square centimetre. The dimensions of pressure are [Force]÷[Area], or $[L^1MT^{-2}]\div[L^2]$, or $[L^{-1}MT^{-2}]$.

If the pressure over a surface is not uniform, then we measure the pressure at a point by considering the force exerted on an element of area, taken round the given point, so small that the pressure is practically constant over this area, and divide the force by the area; a process exactly analogous to that adopted in the case of a variable speed in § 31.

It also follows, as a consequence of the mobility of fluids, that if we apply a pressure to a fluid enclosed in a vessel, then the fluid will transmit this pressure equally in all directions.

If a fluid is unacted upon by any other forces besides the pressure of the sides of the containing vessel, then the pressure must be the same at every point within the fluid, and must act at every point equally in all directions. This statement may be proved by imagining that a small cubical element of volume of the fluid becomes solidified without any other change. This element will evidently still remain in equilibrium, and hence the forces acting on all the faces must be equal. As the area of all the faces is the same, this means that the pressure on all the faces must be the same. Since this must hold good however the small cube is turned, it follows that the pressure must be the same in all directions.

127. Fluids under the Action of Gravity. Principle of Archimedes.—In a fluid at rest, and acted upon by gravity, the pressure in the lower layers is greater than in the upper, since each layer has to support the weight of all the superincumbent layers. The pressure throughout any horizontal layer must, however, be everywhere the same. Otherwise, if there were two points in the same horizontal plane at which the pressure was different, then, since no work would be done against gravity by the passage of fluid from one of these points to the other, if we had a small pipe with one end at one point, and the other end at the other, the fluid would flow from the point of higher pressure to the point of lower pressure through the pipe. This motion would also take place even if no pipe connected the two points, and hence the fluid would not be at rest, which is contrary to hypothesis. If the two points are at different levels, then the pressure at the lower point is greater, but the liquid there does not move to the higher point, since, to do so, work would have to be done against gravity.

When a solid body is immersed either partly or wholly in a fluid, it displaces a volume of the fluid equal to the volume of the immersed part, and it experiences an upward force, due to the fluid, equal in magnitude to the *weight* of the volume of fluid displaced. This is known as the *Principle of Archimedes*, and its truth may be proved as follows. Suppose that the body immersed in the fluid is a cube *abcdefgh* (Fig. 97), and that it is immersed so that the edges *ae*, *bf*, *cg*, and *dh* are vertical. Then the total pressure of the fluid on the face *adhe* is exactly equal and opposite to the total pressure on the face *bcgf*. For we may suppose them each divided into equal horizontal strips, so that the pressure is constant over each strip. Then for each strip in one face there is an equal strip in the other in the same horizontal plane, so that the pressure is the same. Hence the total

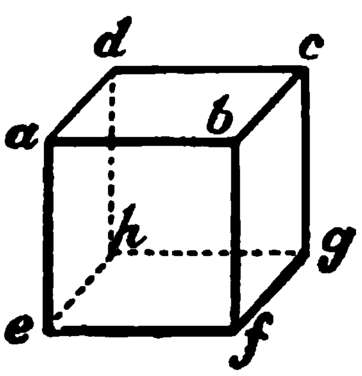


FIG. 97.

pressure exerted on each pair of strips is equal and opposite; and therefore the total pressure on one face is equal to the total pressure on the other. The same argument applies to the faces *abfe* and *dcgh*. The upward pressure on the face *efgh* is, however, greater than the downward pressure on the face *abcd*, since it is at a lower level. In order to see what is the difference between these two forces, suppose the cube removed and replaced by a cube of the given fluid, which by some means has been solidified without any other change. This cube will then be in equilibrium in the fluid. The total pressures on the vertical faces will as before exactly balance each other, so that this cube of the fluid is in equilibrium under the three following forces: (1) the weight of the cube of fluid acting downwards; (2) the total pressure of the fluid on the upper face *abcd* acting downwards, and (3) the total pressure of the fluid on the lower surface *efgh* acting upwards. These forces are all parallel, and so, in order that there may be equilibrium, the sum of the two downward acting forces must be equal to the upward acting force; that is, the difference between the total upward fluid pressure and the total downward fluid pressure is equal to the weight of the cube of fluid. Hence when the solid is immersed in the fluid, since its upper and lower faces occupy just the same positions, as we have supposed the faces of the fluid cube to occupy, the difference in the total pressures[1] on the lower and upper faces will be the same as before, *i.e.* it will be equal to the weight of a cube of the fluid equal in size to the solid, or, in other words, the upward force acting on the immersed solid will be equal to the weight of the fluid displaced by the solid. An experimental proof of the truth of Archimedes's principle will be given later, when considering the properties of liquids.

128. Expansive Power of Gases.—The property of gases which distinguishes them from other fluids is that a given mass of gas, when introduced into a closed vessel, always exactly fills the vessel, whatever its volume. Thus if we have two equal closed vessels connected together by a tube which can be closed by means of a tap, and one of these vessels is filled with a gas, say air at the ordinary pressure, while the other does not contain any matter, or, in other words, has a *vacuum* inside, then, on opening the tap, the air immediately expands and rushes into the second vessel, till finally there is the same quantity of gas in each vessel. By again closing the tap and exhausting the air from one of the vessels by means of an air-pump (§ 136), and then opening the tap, the remaining gas again expands and fills the two vessels. This operation may be indefinitely repeated, and in every case the gas left in the one vessel will, when the tap is opened, expand and fill the two vessels. This experiment illustrates the expansive power of gases.

Since the gas enclosed in a vessel always expands and completely

[1] In many cases where no ambiguity can be caused, we shall use the word pressure where total pressure is meant.

fills the vessel, even if this latter is increased in volume, it follows that the gas must exert a pressure on the inside of the containing vessel. That this is so can be shown by enclosing some air at ordinary atmospheric pressure in a thin glass flask, and then removing the air from outside the flask by placing it beneath the receiver of an air-pump. When, unless the flask is fairly strong, the pressure exerted by the air inside the flask will be sufficient to burst the flask. The reason that the flask does not burst before the air surrounding it is removed, is that this air presses on the outside of the flask and counteracts the effect of the pressure of the enclosed air on the inside. When the air outside is removed by means of the pump there is no pressure exerted on the outside, and the flask may not be strong enough to withstand the inside pressure.

129. Density of Gases.—The density of a body is the mass of unit volume. Hence in the *c.g.s.* system the density of a body is the mass in grams of a cubic centimetre. If the gram were exactly the mass of a cubic centimetre of water at 4° C. (§ 146), the density of water at this temperature would be unity. Since, as will be explained later on, the density of bodies is generally obtained by determining the ratio of their density to that of water, it is usual to assume that the density of water at 4° C. is exactly 1. If it is required to call attention to the fact that the density has been obtained on this assumption, it may be referred to as the relative density.

The term specific gravity is often used as synonymous with what we have called the density, but is sometimes distinguished as the *weight* of unit volume. In the case of gases, the density is sometimes referred to that of hydrogen or air, taken as unity. The following table gives the density of several of the more important gases: (1) in grammes per cubic centimetre; (2) with reference to hydrogen, taken as unity, the gas being in every case at a temperature of 0°C., and under a pressure of one standard atmosphere.

DENSITY OF GASES.

Gas.	Density in Grammes per Cubic Centimetre.	Density relative to Hydrogen taken as Unity.
Hydrogen . . .	0.0000896	1.00
Coal gas . . .	0.00046 (about)	5.1 (about)
Nitrogen . . .	0.001257	14.03
Air	0.001293	14.43
Oxygen	0.001430	15.96
Carbon dioxide . .	0.001974	22.03
Chlorine . . .	0.003133	34.97

130. Elasticity of Gases.—The only kind of elasticity of which a gas is capable is elasticity of volume or bulk elasticity, since it is only to a change of volume that a gas opposes any resistance. If the pressure acting on a volume V of a gas is increased from P to $P+p$, and as a

result the volume becomes reduced to $V-v$, then the strain or deformation produced in a volume V is v, and therefore the strain produced per unit volume is v/V, while the stress corresponding to this strain is p. Hence, since the elasticity of a body is the ratio of the stress to the strain it produces, the elasticity of the given gas is

$$p \div \frac{v}{V} \text{ or } \frac{pV}{v}.$$

In order to study the elasticity of gases, Robert Boyle made use of a glass U-shaped tube (Fig. 98). The end of the shorter limb AB of this tube was closed at A, while the end of the longer limb was open. Having calibrated the shorter limb, so that the volume occupied by the gas enclosed in it was known, mercury was poured into the open limb, so as to imprison a certain quantity of air in the closed limb. The volume of the enclosed gas was obtained from the previous calibration of the tube, and the pressure to which it was subjected was the atmospheric pressure (§ 133), together with the weight of a column of mercury of height ED. By adding more and more mercury, and reading the corresponding values of the volume and the height ED, Boyle obtained a series of values of the volume of a given mass of air under different pressures, and as a result he was led to enunciate a law, which is known as Boyle's[1] law, that the product of the volume of a given mass of a gas into the pressure is a constant for all pressures.

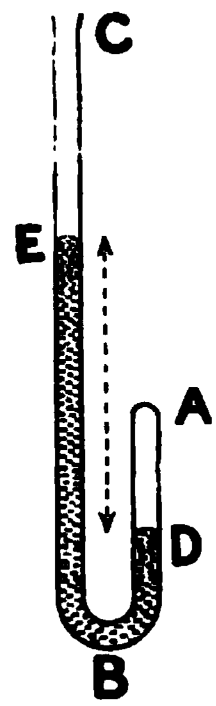


FIG. 98.

If V is the volume of a given mass of gas, and P is the pressure to which it is subjected, then Boyle's law states that, so long as the temperature remains constant,

$$PV = \text{constant}.$$

The experiments of Boyle only showed that the volume of a gas is inversely as the pressure for a small range of pressures, and his form of apparatus was not sufficiently accurate to detect small deviations from his law, even if they had occurred within the range of his experiments.

Regnault in 1847 designed a form of apparatus in which the two great defects of Boyle's apparatus, namely that the possible range of pressures is small, and that as the pressure increases the volume of the gas becomes so small that the inevitable errors made in measuring this volume bear a larger and larger ratio to the volume to be measured, were to a great extent eliminated.

[1] Boyle published the experiments on which he founded the enunciation of his law in 1662. Fourteen years later Mariotte enunciated the same law, using the same curious illustration of a heap of sponges as had been used by Boyle. The law is, however, known on the Continent as Mariotte's law. For an interesting discussion on the true discoverer of this law, see Tait's "Properties of Matter," p. 307.

Regnault used a column of mercury to measure the pressure to which the gas was subjected, but instead of being limited to a column about a metre high, he used one 30 metres high. This column of mercury was contained in a series of strong glass tubes, which were clamped end to end and fixed to the side of a tower. In order to allow of the glass tubes expanding and contracting with changes of temperature without fear of fracture, each tube, instead of being rigidly fixed to the wall, was suspended by two strings which passed over pulleys fixed to the wall and had counterpoises attached (Fig. 99). In this way the weight of the whole column was distributed between these supporting strings, and any expansion could take place quite freely. In order to measure the height of the mercury column a series of fine marks, 0, 1, 2, 3, 4, &c., were engraved on the glass, and the distance between these marks was measured by means of a cathetometer.

Fig. 99.

The bottom tube was firmly cemented at B to a strong iron reservoir LF, into which the glass tube CD to contain the gas under experiment was also cemented. Attached to the top of this reservoir was a small force-pump G, by means of which water could be forced into F from the vessel H, thus driving the mercury which filled the lower part of F up the open tube AB, and compressing the gas in the closed tube DC; the magnitude of the pressure being obtained by observing the height to which the mercury rose in AB. The tube CD was furnished with a tap at C, by means of which it could be closed, or when this was open the gas under experiment could be forced in from a reservoir by means of a force-pump attached to the tube K. A wide tube surrounded CD, so that by the passage of a current of water the temperature of the enclosed gas could be kept constant, as indicated by the thermometers T_1 and T_2. Two fine crosses were engraved, one at D and the other at E, and the volume of the closed tube from the tap down to each of these crosses was very carefully determined.

When making an experiment, the gas was pumped into CD till the

surface of the mercury stood at the mark D in the closed tube, and at the same level in the open tube. The gas was therefore at atmospheric pressure. Water was then pumped into F till the surface of the mercury stood at E, and the position of the top of the mercury column in AB, measured from E, was read, and thus the new pressure was obtained. Then, the pressure being kept constant, gas was pumped in through K till the surface of the mercury was at D. More water was then pumped into F, till the gas was compressed to the volume CE, and the pressure noted as before. More gas was then pumped in, and the series of operations repeated till the greatest available pressure was reached.

From the readings thus taken it could be seen what increase of pressure was necessary to compress the gas from the volume CD to the volume CE, starting at different initial pressures; the great improvement on the previous methods being that the initial and final volumes were the same at the high pressures as at the low, and hence the inevitable small uncertainties made in measuring the volume did not produce a greater percentage error at the high pressures than at the low.

In deducing the pressure from the height of the mercury column, Regnault allowed for the change in density of the mercury with temperature, the temperature of the column being measured by a series of thermometers T_3, T_4, &c., placed alongside the column. He also allowed for the increase in density of the mercury at the lower parts of the column produced by the pressure of the superincumbent mercury. Finally, since to obtain the total pressure to which the gas is subjected we must add the pressure of the atmosphere (§ 133) on the top of the mercury column in the open tube, he read the height of a barometer placed on a level with the surface of the mercury in AB for each position.

In the following table some of Regnault's results are given. The first column contains the initial pressure (P_0) in centimetres of mercury under which the gas occupied the volume CD, which may be called V_0. If then P_1 is the pressure when the volume is reduced to CE, say V_1, then, if Boyle's law is exactly true, V_0P_0 would be equal to V_1P_1, or the ratio $\frac{V_0P_0}{V_1P_1}$ would be unity; the actual values found for this ratio are given in the second column of the table:—

Air.		Carbon Dioxide.		Hydrogen.	
P_0.	$\frac{V_0P_0}{V_1P_1}$.	P_0.	$\frac{V_0P_0}{V_1P_1}$.	P_0.	$\frac{V_0P_0}{V_1P_1}$.
Cm. of Mercury.		Cm. of Mercury.		Cm. of Mercury.	
73.872	1.001414	76.403	1.007597	221.118	0.998584
211.253	1.002765	318.613	1.028698	584.518	0.996121
414.082	1.003253	487.977	1.045625	917.650	0.992933
933.641	1.006366	961.997	1.155865		

From this table it will be seen that in the case of air and carbon dioxide V_0P_0 is always slightly greater than V_1P_1, and that as the pressure increases this excess becomes greater and greater. Hence these gases are slightly more compressible, particularly at high pressures, than they would be if they obeyed Boyle's law exactly. Hydrogen, however, deviates from Boyle's law in the opposite direction, V_0P_0 being *less* than V_1P_1, so that this gas is less compressible than a gas which obeys Boyle's law exactly.

Regnault in his experiments was only able to go up to a pressure of 27 atmospheres (§ 133). Amagat has, however, investigated the

FIG. 100.

elasticity of gases up to pressures of about 300 atmospheres, and his results for hydrogen, nitrogen, and carbon dioxide (at two temperatures) are shown in Fig. 100. In this figure a curve has been drawn representing for each gas the values of the product PV at different pressures. Since, if a gas obeys Boyle's law exactly, PV is constant, the curve

corresponding to such a gas would be a horizontal straight line parallel to the axis pressures.

It will be noticed that the curve for hydrogen slopes upwards for increasing pressures, indicating that the gas is less compressible, *i.e.* more elastic, than if it obeyed Boyle's law. In the case of nitrogen at pressures below 40 metres of mercury the curve slopes downwards, and the gas is less elastic than if it obeyed Boyle's law; while for higher pressures it resembles hydrogen, in that its elasticity is greater than that given by Boyle's law. Carbon dioxide at a temperature of 100° C. gives a curve which is an exaggeration of the nitrogen curve. At a temperature of 35°.1 the curve for carbon dioxide has a very distinctive form, there being a pressure (70 metres of mercury), for which the product VP has a sharply marked minimum value.

A consideration of these curves shows that gases, which at low pressures deviate from Boyle's law in that they are too compressible, at high pressures and temperatures resemble hydrogen at ordinary pressures, and deviate from Boyle's law in the opposite sense to that at low pressures.

131. The Air Manometer.—The elasticity of a gas can be made use of to measure pressures. An instrument for this purpose consists of a curved tube ABC (Fig. 101) closed at one end, A, with some mercury in the bend enclosing some air in the closed limb, the volume of which can be read off on a scale attached to the side of the tube. The open end C being connected with the vessel in which the pressure has to be measured, suppose the volume of the air to be reduced from V_0, at atmospheric pressure, to V_1, the mercury in the tube standing at E and D in the two branches of the tube. Then the pressure acting through C is balanced by the elasticity of the air, together with the weight of a mercury column of height $\overline{\text{DE}}$. The pressure due to the elasticity of the air is by Boyle's law equal to

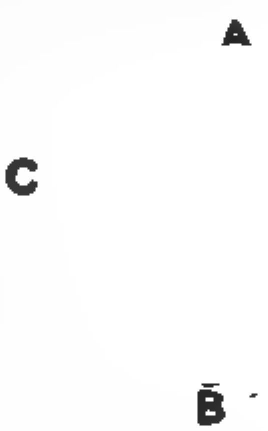


FIG. 101.

$\frac{V_1}{V_0}$ atmospheres, and hence the pressure to be measured is equal to $\frac{V_1}{V_0}$ atmospheres together with the weight of the column of mercury $\overline{\text{DE}}$. Of course for high pressures a correction would have to be applied, to allow for the deviation of air from Boyle's law.

132. Torricelli's Experiment.—In the year 1643, an Italian named Torricelli having filled a glass tube, about a metre long and closed at one end, with mercury, inverted it and dipped the open end below the surface of some mercury in a trough. He then found that, instead of continuing to completely fill the tube, the mercury forsook the upper part of the tube, the height of the column being about 76 centimetres. Torricelli gave the true explanation of this phenomenon, namely, that the mercury column was supported by the pressure of the atmosphere

acting on the free surface of the mercury in the trough, so that this pressure was equal to the weight of a column of mercury about 76 cm. high. This explanation also accounted for the elevation of water in suction-pumps, which had previously been *explained* by saying that nature abhorred a vacuum, and that as the plunger of the pump rose, it tended to produce a vacuum, and therefore the water rushed in. Torricelli's experiment was further extended by Pascal, who tried the experiment with tubes filled with oil, water, and wine, the height of the column being in each case inversely proportional to the density of the liquid employed. Pascal also suggested that if Torricelli's explanation were the correct one, then the maximum height of the mercury column, or the height of the *barometer*, as it is called, would be less at the top of a mountain than at the foot, since the air is a heavy fluid, and therefore the pressure increases with the depth. This experiment was carried out, and the results completely confirmed his prediction.

133. Pressure of the Atmosphere.—Since the volume of a gas is so very largely affected by the pressure to which it is subjected, it is necessary to state the pressure at which the measurement is made. To simply state that the measurement was made at "atmospheric pressure" is, in many cases, not accurate enough, for it is found that the barometric height, and therefore the pressure of the atmosphere, varies by a considerable amount from time to time. A standard pressure has therefore been adopted, which is called *an atmosphere*, or simply *the standard pressure*. This pressure is such that it would support a column of mercury 76 cm. high. Since the density of mercury varies with the temperature, and the pressure necessary to support a given height depends on the density of the mercury, it is necessary to state the temperature of the mercury when defining the standard pressure. In addition, since the pressure necessary to support the column of mercury depends on the *weight* of the mercury, and the weight of a column of mercury of given height depends on the value of g, or the acceleration due to gravity, it is necessary to state the value of g, for which 76 cm. of mercury are equal to the standard atmosphere. The temperature chosen has been that of melting ice (0° C.), and the value of g, since g varies both with the latitude (§ 116) and with the altitude, is taken as that at latitude 45° and at the sea-level. Hence the standard pressure is defined as such that it will support a column of mercury 76 cm. high, at latitude 45° and at the sea-level, the temperature of the mercury being 0° C. The density of mercury at 0° C. is 13.596, and the value of g at the sea-level and at latitude 45° is 980.60 cm. per sec. per sec. Hence the value of the standard pressure is

$$76 \times 13.596 \times 980.60 \text{ dynes per sq. cm.}$$
$$= 1013250 \text{ dynes per sq. cm.}$$

This number is very nearly one million dynes per square centimetre,

and it has been proposed to take a pressure of exactly one million or 16^6 dynes, or a mega-dyne, per square centimetre as the standard pressure. This would correspond to a column of mercury at 0° C., at latitude 45° and the sea-level, of a height of

$$\frac{1000000}{13.596 \times 980.60} = 75.005 \text{ cm.}$$

134. The Barometer.—A knowledge of the pressure exerted by the atmosphere, or the height of the barometer, is of great importance not only in meteorology, but also, as we shall see in the later sections of this volume, in many branches of physics. An instrument designed for measuring the pressure of the atmosphere is called a barometer, and we shall now proceed to describe one or two of the more important kinds of barometers.

Barometers may be divided into two classes: (1) Liquid barometers, in which the pressure is measured in terms of the height of a column of a liquid, and (2) aneroid barometers, in which the pressure is measured by the deformation of the lid of a metal box.

Practically the only liquid that is used in liquid barometers is mercury, since, on account of its great density, the height of the column which the pressure of the atmosphere can support is of a manageable magnitude. Another advantage possessed by mercury is that it does not absorb moisture from the air, as does glycerine, the only other liquid that has been used to any extent.

The simplest form of mercury barometer is the syphon barometer. It consists of a U-shaped tube, the longer limb (AB, Fig. 102) of which is closed at A, and is about 86 cm. long, while the shorter limb is open at C. This tube is filled with pure mercury, and by boiling the mercury any air or moisture adhering to the mercury or the bore of the tube is expelled. The distance $\overline{DE}$ is equal to the barometric height. When the barometric pressure increases, the mercury rises in the closed limb and falls in the open limb; and if the bore of the two limbs is the same, the movement of the mercury surface (meniscus) is the same in the two limbs but in opposite directions. Hence, if the mercury rises in the closed limb by 1 cm. it will also fall in the open limb by 1 cm., and therefore the distance $\overline{DE}$ will increase by 2 cm., that is, the atmospheric pressure will have increased by two centimetres of mercury.

FIG. 102.

If a scale is attached to either of the tubes, and each half-centimetre is marked a centimetre, then the reading at once gives the height of the barometer. Since, however, the bore of a glass tube is never quite uniform, two scales are fixed, one alongside each limb, having their zeros on the same horizontal plane, that alongside the closed limb reading upwards, and

that alongside the open limb reading downwards. The sum of the readings corresponding to the two mercury surfaces then gives the height of the mercury column.

In the cistern barometer the tube is straight, the open end dipping below the surface of some mercury contained in a fairly wide vessel. Since, as the atmospheric pressure alters, and therefore the height of the mercury column alters, mercury either enters or leaves the tube, the level of the mercury in the cistern will alter. As it would be inconvenient to have a cistern with such a large cross section, in proportion to that of the tube, that such fluctuations in the quantity of mercury contained in the tube as occur in practice should not *appreciably* alter the level of the surface in the cistern, a device due to Fortin is employed, by means of which the level of the surface of the mercury in the cistern is always brought back to a fixed mark connected to the scale by which the height of the column is measured. The plan adopted is shown in Fig. 103, and consists in making the bottom of the cistern flexible. The upper part of the cistern is of glass, and is cemented to a boxwood ring A, to which is tied a ring of buckskin B. This buckskin carries a wooden button C, which rests on the point of a screw S, and forms a flexible bottom to the cylinder, so that the surface of the mercury in the cistern can be raised or lowered by turning the screw. A small, pointed ivory pin, D, is fixed to the top of the cistern, and the graduations of the scale, which are usually engraved on a metal tube surrounding the glass barometer tube, count from the point of this pin. Before making a reading of the meniscus of the mercury in the tube, the surface of the mercury is adjusted till it exactly touches the point of the ivory pin. This adjustment is complete when the point of the pin appears just to touch its image, as seen by reflection in the surface of the mercury.

FIG. 103.

FIG. 104.

The aneroid barometer consists essentially of a cylindrical metal box A (Fig. 104), the lid of which consists of a thin corrugated

metal plate. The inside of this box is exhausted by means of a pump, leaving a more or less perfect vacuum, and the pressure of the air, acting on the thin elastic lid, bends it and forces it in to a certain extent. As the pressure of the atmosphere varies, the amount of flexure of the lid varies, and by means of a system of delicate levers, C, D, E, this change in the flexure of the lid is shown by the movement of a pointer, F, over a graduated scale. The great advantage of an aneroid barometer over a mercurial barometer is its extreme portability. The scale of all aneroids, however, has to be set out by comparing them with a mercurial barometer.

135. Corrections to Barometer Reading.—In addition to the corrections to reduce the height of the mercury column to what it would be at 0° C., at latitude 45°, and at the sea-level, a correction has to be applied to allow for the expansion of the scale by means of which the height of the column is measured. If this scale is correct at 0° C., then at all temperatures above 0° C., the length of the divisions will be too *great*, since all metals increase in length when heated. Let a be the coefficient of linear expansion of the metal of which the scale is composed (§ 184), so that unit length of the scale at 0° C. becomes $1+a$ at 1°C. and $1+at$ at t°C. If h_t is the barometric height as measured with the scale at a temperature t, then the height as measured with the scale at 0° would be greater, since the length of each division of the scale would be less in the ratio of 1 to $1+at$, so that the number of the divisions corresponding to a given length (*i.e.* the length of the mercury column) will be increased in the ratio of $1+at$ to 1. Hence if h_0 is the barometer reading corrected for the expansion of the scale,

$$h_0 = h_t\,(1+a_t).$$

Now h_0 is the height of a column of mercury at a temperature t, and we have to find what would be the height if the temperature of the mercury were 0° C. If d_t is the density of the mercury at $t°$, d_0 the density at 0°, δ the coefficient of cubical expansion of mercury (see § 189), and H the height which the column would have if the mercury were at 0° C.; then 1 c.c. of mercury at 0° becomes $1+\delta$ c.c. at 1°, and $1+\delta t$ c.c. at $t°$. Hence, since the mass M of the mercury remains the same,

$$M = V_0 d_0 = V_t d_t,$$

where V_0 and V_t are the volumes of the mass M at the temperatures 0° and $t°$ respectively. Therefore

$$d_0 = (1+\delta t)d_t,$$

$$\text{or}\quad \frac{d_t}{d_0} = \frac{1}{1+\delta t} = 1 - \delta t + \delta^2 t^2 +, \text{ \&c.}$$

Since δ is an excessively small quantity, we may neglect the term involving δ^2 and higher powers of δ. Therefore

$$\frac{d_t}{d_0} = 1 - \delta t.$$

The height of a column of liquid supported by a given pressure being inversely proportional to the pressure,

$$\frac{H}{h_0}=\frac{d_t}{d_0}=1-\delta t.$$

Hence

$$\begin{aligned} H&=h_0(1-\delta t)\\ &=h_t(1+at)(1-\delta t)\\ &=h_t(1-(\delta-a)t), \end{aligned}$$

if we neglect the term $\delta a t^2$, which is excessively small.

For brass $a=.000020$, and for mercury $\delta=.000182$.

Hence, for a mercury barometer with a brass scale, the *reduced* height corresponding to an observed height h_t, at a temperature t, is given by

$$H=h_t(1-0.000162t).$$

This height H corresponds to a pressure of Hg dynes, where g is the acceleration of gravity at the place of observation. If g_{45} is the value of g at latitude 45°, and at the sea-level,

$$\frac{g}{g_{45}}=1-0.0026 \cos. 2\lambda-0.0000002l,$$

where λ is the latitude of the place of observation and l is the height above the sea-level in metres.

If H_0 is the height, under standard conditions, which corresponds to the same pressure as does H at the place of observation, then

$$Hg=H_0g_{45},$$

$$\text{or}\quad H_0=\frac{Hg}{g_{45}},$$

$$=h_t(1-0.000162t)(1-0.0026 \cos. 2\lambda-0.0000002l).$$

The above-mentioned corrections are practically common to all mercury barometers, since the scale is almost invariably made of brass, and the magnitude of the corrections is the same for all barometers under the same conditions. There is, however, a correction which depends on the fact that the surface of the mercury in the tube (meniscus) is curved and not plane. Hence, on account of capillarity (see § 160), there is a force tending to depress the mercury column, and on this account the height of the column is less than it would be if the atmospheric pressure were counterbalanced by the weight of the column alone. The amount of the correction to be applied to allow for this capillary depression of the column depends on the diameter of the bore of the tube, and for tubes of which the diameter exceeds 2.5 cm. it can be entirely neglected. The corrections to be applied to barometers having smaller bores are given below, but it must be remembered that these corrections are only approximate, and the only satisfactory method of

finding the capillary correction for a barometer is to compare its reading with that of a standard barometer of which the bore is more than 2.5 cm. in diameter.

Diameter of Bore . . .	0.4 cm.	0.8 cm.	1.2 cm.
Capillary Depression . .	0.14 cm.	0.05 cm.	0.02 cm.

In every case the capillary correction must be added to the observed height, since capillarity tends to give too low a value for the barometric height. In the syphon barometer the effects of capillarity in the two limbs tend to neutralise each other; but since in one case the mercury surface is quite clean and exposed to a vacuum, while in the other case it becomes coated with dust and is exposed to air, this compensation is by no means complete.

136. Mechanical Air-Pump.—An air-pump is an instrument for withdrawing the air from within a vessel. In its simplest form the air-pump consists of a cylinder in which a piston P (Fig. 105) fits air-tight. There is a hole through the piston closed by a flap valve C, which can open outwards. A pipe, the opening to which can be closed by a valve, B, which opens inwards, leads to the vessel D, that is to be exhausted. When the piston is drawn upwards the valve C closes, and the pressure below the piston is reduced so that the air in the receiver, on account of its elasticity, is able to raise the valve B, and flows into the cylinder. When the piston descends the valve B closes, and the air in the cylinder is compressed till it is able to force open the valve C, and escape into the air. By repeating this process the air is gradually pumped out of D.

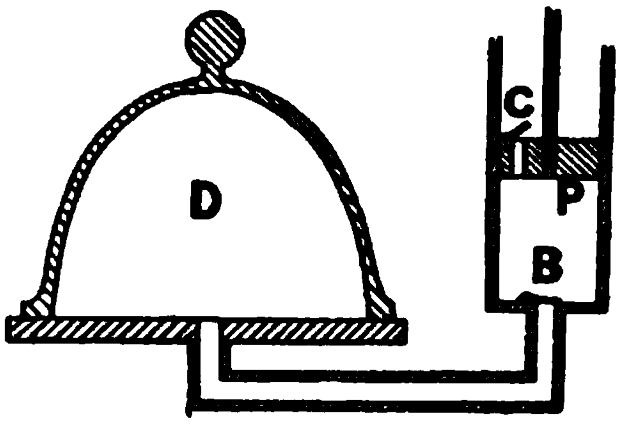


FIG. 105.

If the volume of the vessel D and the pipe connecting it to the cylinder is V, and the volume of that part of the cylinder through which the lower surface of the piston moves during a stroke is v. Then, if we start with the piston at the bottom of its stroke, the volume of the mass (m) of air in the instrument is V. At the end of the upward stroke the volume of this mass of air will be $V+v$. Of this volume v c.c. will be expelled at the down-stroke, and V c.c. will be left in the instrument. Hence at the end of the first stroke the mass of air in the receiver is $\frac{V}{V+v}m$. At the end of the second up-stroke the volume of this mass of air expands to $V+v$, and during the down-stroke v c.c. of air at this density are expelled. Hence the mass of air at the end of the second stroke left in the receiver is $\frac{V}{V+v}$ of the mass of air in the receiver

before the second stroke, or $\frac{V}{V+v} \cdot \frac{V}{V+v}\, m$, which may be written $\left(\frac{V}{V+v}\right)^2 m$.* In the same way the mass of air left after three strokes is $\left(\frac{V}{V+v}\right)^3 m$, and generally the mass of air left after n strokes is $\left(\frac{V}{V+v}\right)^n m$. Since this mass of gas now occupies V c.c., it follows that its density is $\left(\frac{V}{V+v}\right)^n m/V$, while the original density was m/V. Hence the density of the air left in the receiver after n strokes is $\left(\frac{V}{V+v}\right)^n$th. of what it was originally. From this expression it will be seen that we cannot make the density of the air zero unless the number of strokes n is infinite. If the original pressure within the receiver is p, then after n strokes it will be $\left(\frac{V}{V+v}\right)^n p$. In practice, however, it is not possible to obtain a very low pressure on account of mechanical defects in the construction of the valves, to leakage round the piston, and to the fact that the piston, when at the bottom of its stroke, does not completely expel all the air in the cylinder. Another difficulty met with in mechanical pumps is that, after the exhaustion has proceeded a certain distance, the elasticity of the air left in the receiver is not great enough to lift up the valve B (Fig. 105), so that the air left in the receiver cannot escape into the cylinder. In order to overcome this difficulty, the valve is often carried at the end of a rod A (Fig. 106), which passes through the piston with a little friction. When the piston starts moving up, it raises the valve A as far as a stop fixed to the top of the rod will allow. When the piston commences to descend it forces the valve down into its seating, and thus closes the connection between the cylinder and the receiver. In this way the valve is opened by the movement of the piston, and not by the elasticity of the air in the receiver.

FIG. 106.

In the Fleuss pump, which is shown diagrammatically in Fig. 107, the difficulty with the valves, and also the defect that in the old form of pump there is always a certain amount of clearance between the bottom of the piston and the cylinder, so that all the air contained within the cylinder is not expelled during the downward stroke, is avoided in another way. The piston rod H passes air-tight through a partition C, in which

FIG. 107.

is a valve that opens upwards. The lower part of the cylinder is filled with oil, E, and has two side-tubes, A and B, which are connected to a small vessel, F, that serves to prevent the oil being splashed up the tube, G, which is attached to the vessel being exhausted.

There is a certain amount of oil above the piston, so that when the piston is at the top of its stroke, this oil, flowing up through the valve C, chases out all the air. As the piston is forced down, a vacuum is produced above, since the valve C closes, and when it reaches the bottom of its stroke, it has passed below the side-tube A, and thus the air rushes in from the vessel being exhausted, and fills the barrel. When the piston rises above the side-tube A, the connection between the air enclosed in the cylinder and the vessel being exhausted is cut off, while the air within the cylinder is, during the up-stroke, forced out through C. The object of the double side-tube is to allow the oil, at the bottom of the stroke, to flow round from below the piston to above.

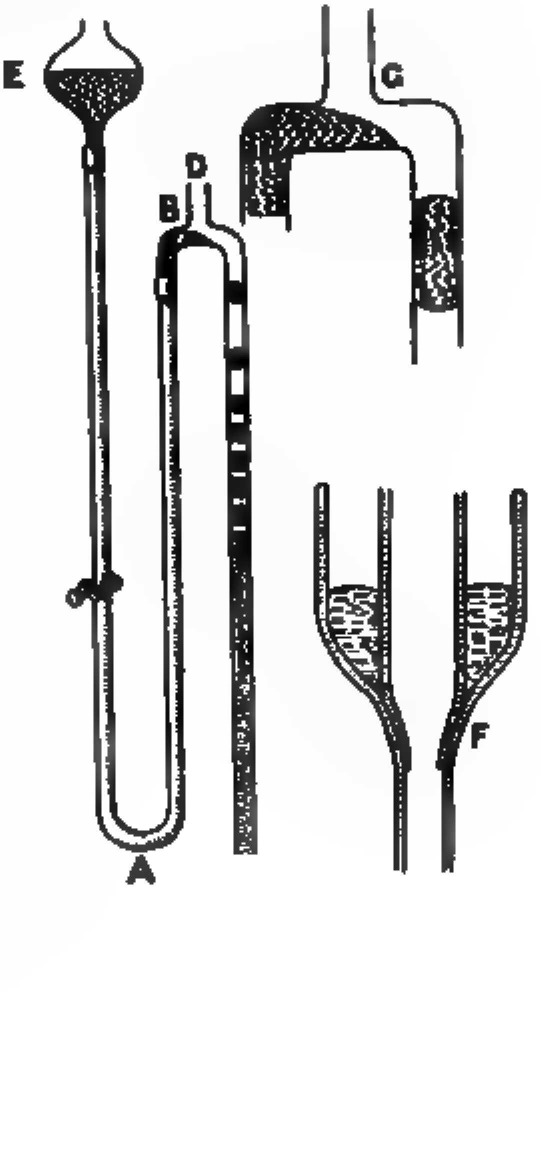


Fig. 108.

137. Mercury Air-Pumps.—A very good mechanical pump will exhaust a vessel till the pressure of the remaining air will support a column of mercury of about 0.05 millimetre in height. In order to get a better vacuum than this, it is necessary to employ a pump in which the piston is formed by a quantity of mercury. Sprengel's mercury-pump consists of a bent glass tube ABC (Fig. 108), with a side-tube D joined on at the bend. The end A of this tube is connected by means of a thick-walled rubber tube with a reservoir E containing mercury. The vessel to be exhausted is connected to the side-tube D, generally by means of a glass tube fitting by a well-ground neck into the end of the tube D. This ground joint is surrounded by a glass cup, shown on a large scale at F, in which a little mercury is placed to prevent the external air reaching the joint. The flow of mercury from the reservoir E is adjusted by a pinchcock on the rubber tube, so that when the mercury reaches the top of the bent tube (shown on a larger scale at G) it does not pass over in a continuous

stream, but breaks up into drops. Each drop, as it falls down the tube BC, forms a small piston which carries a little of the air from the space H before it, the air in the vessel attached to D expanding to take its place. These small mercury pistons carry the air down the tube, and finally drive it out at the open end C, the mercury being caught in a vessel K, and returned every now and then to the reservoir E. When the exhaustion is getting fairly complete, each mercury pellet, as it falls down the tube BC, strikes the top of the mercury column left in the tube, which has a height practically equal to the barometric height, with a sharp metallic click. When the exhaustion is not so complete, the air imprisoned between each pellet of mercury acts as a buffer, and there is no click. The object in carrying the tube from the reservoir E to B down to A is that, supposing the supply of mercury in E runs short, the mercury in the connecting tube EAB will not be driven over by the atmospheric pressure, and thus admit the air to the vessel being exhausted, but will simply stand so that the difference in level between the surface of the mercury in the part of the tube EA and that in the part AB will be equal to the barometric height, and this column of mercury will prevent the access of the air.

Another form of mercury-pump is shown in Fig. 109, and was devised by Töpler. A cylindrical glass vessel A has a side-tube B attached, and to the lower end of this side-tube is attached another tube CH, which is connected to the vessel to be exhausted. To the top of A a tube DG, about 80 cm. long, is attached, while to the bottom another tube EF, also about 80 cm. long, is attached. The lower end of EF is connected by a length of rubber tubing with a reservoir K containing mercury. When K is raised the mercury rises in FE, and when the surface reaches C cuts off the connection between the vessel A and the tube CH. K is raised till the mercury completely fills A and flows out through G, driving any air that was in A before it. If now K is lowered, so that the surface of the mercury in K is more than 76 cm. below C, the mercury will fall in A and in DG till it stands at the barometric height in DG, and will leave a Torricellian vacuum in A. When the mercury in EF has fallen below C, the tube HC will be connected to this vacuum, and hence the air in the tube HC and any vessel attached to H will rush into A. By again raising K the air enclosed in A will first be cut off from CH by the rising mercury and then forced out of the apparatus at G, and on lowering K a vacuum will again be left in A.

Fig. 109.

The mercury here plays the part of a piston moving up and down in the cylinder A, and constitutes its own valves. By repeating this operation a number of times it is possible to obtain a very good vacuum, in fact the pressure inside a vessel has been reduced to 0.000012 millimetres of mercury, that is, to .000000016 or .016 of a millionth of an atmosphere.

138. Effusion of Gases.—Suppose a gas of density d (d is the mass in grams of a cubic centimetre, not the density as compared with that of hydrogen) is enclosed in a vessel at a pressure of p dynes per square centimetre above that of the surrounding air, and is allowed to escape through a small opening, the cross section of which is a. Then if v is the velocity with which the gas escapes, *i.e.* the velocity with which a small dust mote would be carried through the opening, the volume of gas which escapes in the unit time is av, and its mass is avd. The kinetic energy of this moving mass of gas is $\frac{1}{2}\, avd.v^2$.

Suppose that a cylindrical tube AB (Fig. 110) of cross section a were fitted over the opening in the vessel, and a small weightless piston C were fitted in this tube and moved without friction, so that as the gas escaped this piston was driven back. If the piston starts in the position C, then at the end of a second it will have arrived at a position D, such that the distance $\overline{CD}$ is equal to v. The pressure acting on the inside surface of the piston exceeds the pressure acting on the outside by ap. Hence the work done by this excess pressure while the piston moves from C to D is $ap \times \overline{CD}$ or apv. This work will be done by the expanding gas as it escapes, whether we imagine such an arrangement of cylinder and weightless, frictionless piston to exist, or whether the gas simply escapes into the air without any such contrivance. It is owing to the expenditure of this work (which is done at the expense of the potential energy of the compressed gas) that the escaping gas possesses kinetic energy, and the amount of the kinetic energy acquired is numerically equal to the work done. Hence

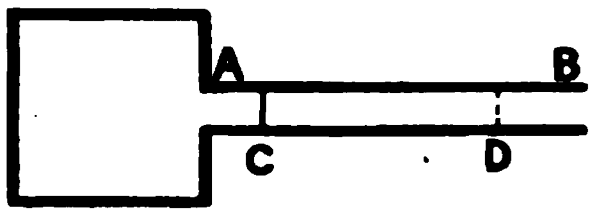


Fig. 110.

$$\tfrac{1}{2}av^3d = apv,$$

or

$$v = \sqrt{\frac{2p}{d}}.$$

This equation gives the algebraic statement of the laws which regulate the rate at which gases effuse through a small opening. These laws were first discovered experimentally by Graham, and are expressed in words as follows : "The velocity with which a gas effuses varies directly as the square root of the difference of pressure on the two sides of the opening, and inversely as the square root of the density of the gas."

Hence it follows that if two gases are allowed to effuse through the

same opening under the same difference of pressure, then the velocities of effusion of the gases will be inversely to one another as the square roots of their densities, or, what amounts to the same thing, the densities of the two gases will be to one another as the square of the times in which equal volumes of the two gases escape through the same opening under the same excess of pressure.

This property of the effusion of gases has been utilised by Bunsen to compare the densities of gases. The instrument he used for the purpose is shown in Fig. 111. It consists of a cylindrical glass vessel A, closed at one end by a stopper S. At V a thin platinum partition completely closes the neck of the tube, but is pierced with a small hole about 0.013 mm. in diameter. The gas under observation is enclosed in A, which is inserted in another vessel containing mercury to such a depth that the top, *r*, of a glass float is on a level with the outside surface of the mercury. The stopper S is then removed, and the time noted that it takes for sufficient gas to escape through the hole in the platinum plate to allow the float to rise till a mark at *t* is level with the surface of the mercury. Then the times obtained with different gases are one to another as the squares of their densities.

FIG. 111.
(From Ganot's "Physics.")

The difference in the rates of effusion of two gases is sometimes used to separate them when they occur as a mixture. Thus by passing atmospheric air through a number of clay tobacco-pipe stems, placed one after the other, a vacuum being maintained at the outside of the tubes, Rayleigh and Ramsey were able to separate argon from nitrogen. The density of argon being about 20 and that of nitrogen 14, we have—

$$\frac{v\ (\text{nitrogen})}{v\ (\text{argon})} = \sqrt{\frac{20}{14}} = 1.2,$$

so that the nitrogen passes through the clay 1.2 times faster than the argon, and the gas which escapes from the end of the pipe-stems is richer in argon than ordinary air. This method of separating gases is called atmolysis.

139. Diffusion of Gases.—If two equal bottles, one containing hydrogen and the other carbon dioxide, are placed mouth to mouth, the

bottle containing the hydrogen being on the top, then after a certain time it will be found that half of the hydrogen has travelled down into the lower bottle and half of the carbon dioxide has ascended into the upper bottle, and this in spite of the fact that the density of the carbon dioxide is twenty-two times as great as that of the hydrogen. This phenomenon is called the diffusion of gases, and we notice that after diffusion is complete the proportion of each gas in each of the two bottles is the same as it would be had one of the bottles filled with either of the gases been connected to a second equal bottle which had been evacuated. Hence it is usual to say that one gas diffuses into a space which is filled with another gas just as if this other gas did not exist. The only effect of the presence of the second gas is that the diffusion, instead of being almost instantaneous, as it would be in the case of a vacuum, takes some time to become complete. It also follows that in the case of a vessel filled with a mixture of gases, each gas exerts the same pressure as it would exert if it were alone present, so that the total pressure is the sum of the pressures exerted by the two gases (Dalton's law). Of course, in each of the above cases it is presupposed that the gases do not exert any chemical action on one another.

It is on account of diffusion that the proportions of nitrogen and oxygen in the air are the same at all elevations, there being no excess of the heavier constituent (oxygen) at low levels or of the lighter constituent high up.

If the two bottles containing hydrogen and carbon dioxide, instead of being placed with their open mouths in contact, are separated by a plate which is pierced with a number of small holes, such as a piece of unglazed porcelain, then diffusion will still go on. First consider the hydrogen; since the bottle containing the carbon dioxide acts as a vacuum, the hydrogen will pass through the pores, and the velocity with which it passes will be given by the same expression as that found for the effusion of a gas in the previous section. Hence the rate at which the hydrogen passes is directly proportional to the square root of the partial pressure[1] exerted by the hydrogen in the hydrogen bottle, and inversely proportional to the square root of the density of hydrogen. In the same way, the rate at which the carbon dioxide passes through the porous plate will be directly proportional to the square root of the partial pressure exerted by the carbon dioxide, and inversely proportional to the square root of the density of the carbon dioxide. At the commencement the partial pressure due to each gas in its own bottle is equal to the atmospheric pressure. Hence the rate at which the hydrogen starts diffusing into the carbon dioxide is to the rate at which the carbon dioxide diffuses into the hydrogen as the square root of the density of carbon dioxide is to the square root of the density of hydrogen; or if v_1

[1] By partial pressure is meant the pressure which would be exerted by the hydrogen alone, suppose the carbon dioxide which is mixed with the hydrogen were removed.

and d_1 are the rate of diffusion of the hydrogen at the start and its density, and v_2 and d_2 are the corresponding quantities for the carbon dioxide, then

$$\frac{v_1}{v_2} = \sqrt{\frac{d_2}{d_1}}.$$

Hence, as the density of carbon dioxide (d_2) is greater than that of hydrogen (d_1), the hydrogen will diffuse more quickly into the bottle containing carbon dioxide than this gas can diffuse into the hydrogen bottle. As a result, the total pressure in the carbon dioxide bottle will become greater than one atmosphere, while that in the hydrogen bottle will be less. The actual pressure causing diffusion being the difference between the partial pressures of the given gas in the two bottles, the acting pressure in the case of the hydrogen will decrease more quickly than in the case of the carbon dioxide, and on this account the rate of diffusion of the hydrogen will decrease more rapidly than that of the carbon dioxide. When, if the two bottles are of equal volume, half of each gas has passed over into the other bottle, the partial pressures exerted by each gas on each side of the porous partition will be equal, and hence diffusion will cease.

140. Absorption of Gases—Occlusion.—Liquids are able to dissolve gases even when they do not enter into any chemical combination with them. The *volume* of gas which the unit of volume of any given liquid can absorb depends on the nature of the gas and on the temperature of the liquid. The number of units of volume of a gas which can be absorbed by unit volume of a given liquid at 15° C. is called the absorption coefficient of the liquid. The absorption coefficients for some gases in water are as follows :—

Ammonia . . .	756	Chlorine . . .	2.4
Carbon dioxide . .	1.0	Oxygen . . .	0.035
Hydrogen . . .	0.019	Nitrogen . . .	0.017

Since a liquid can absorb a certain *volume* of a given gas, it follows that the mass of the gas absorbed by a given volume of liquid depends on the pressure to which the gas and liquid are subjected ; for the mass of a given volume of a gas is, according to Boyle's law, proportional to the pressure, while, as will be seen later, the volume of a given mass of liquid is almost independent of the pressure. Thus at two atmospheres pressure unit volume of water will absorb twice the mass of carbon dioxide that it will at one atmosphere's pressure, at three atmospheres three times the mass, and so on. If, after the liquid has absorbed all it can of a given gas at a given pressure, the pressure is reduced, the excess gas, above the quantity which would be absorbed at this new pressure, will be evolved. This is what happens in the case of soda-water, bottled beer, champagne, &c. In each case the liquid has

absorbed carbon dioxide gas at a high pressure, and when the bottle is opened the excess gas is evolved, and gives rise to the so-called sparkle of the liquid.

Metals such as silver and gold are capable, when in the molten condition, of absorbing gas from the air, just as other liquids ; this gas being evolved when the metal solidifies.

Some metals, notably palladium, are able to absorb very large volumes of hydrogen, even when in the solid state. Thus if a slip of palladium is used as the negative pole in the electrolysis (§ 539) of a dilute solution of sulphuric acid, it will absorb about 900 times its own volume of hydrogen gas. Gases absorbed by solids are said to be occluded.

Almost all solid bodies possess the power of condensing gases on their surface, so that after being surrounded for some time by a gas, a solid becomes coated on the outside with a layer of this condensed gas, which cannot be immediately removed by merely placing the solid under the receiver of an air-pump and producing a vacuum. In order to completely remove this gaseous coating, it is necessary to heat the solid while it is in a vacuum, or to rub the surface with alcohol, or some fine powder, such as tripoli. The amount of gas which can in this way be occluded depending on the surface of the solid exposed, very porous bodies, such as wood-charcoal and platinum-black (*i.e.* finely divided platinum obtained by heating platinic chloride), in which the surface bears a very large ratio to the mass of the body, are able to occlude comparatively large quantities of some gases. Thus freshly heated (in order to free it of occluded air) box-charcoal will occlude about ninety times its volume of ammonia gas.

141*. Kinetic Theory of Gases.—The phenomena of diffusion, in which a heavy gas will move upwards and mix with a lighter gas placed above, and this lighter gas will move down, show that, although such amounts of the gas as we are able to see, and particles of dust suspended in the gas appear at rest, yet there must be some kind of movement going on continuously within a mass of gas. We have seen that the most probable theory of the constitution of matter is to suppose it built up of fine particles or molecules. The kinetic theory of gases supposes that in a gas these molecules are endowed with a motion of translation, and that the spaces between adjacent molecules are fairly great compared with the size of the molecules. We may, as a first approximation, consider that the molecules are hard, elastic spheres, each molecule having a definite mass, and that a gas consists of an enormous number of these small spheres moving about in all directions with different velocities. During its movement each molecule will occasionally collide with another molecule, the two rebounding after collision like two billiard-balls ; also some of the molecules will be continually striking the walls of the vessel containing the gas, and rebounding from them. In the intervals between its impacts with other molecules, or with the

walls, each molecule will travel in a straight line, so that the path of a molecule consists of a zigzag line. On account of their frequent collisions, the velocities of the different molecules must vary considerably, as also the velocity of any given molecule at different times. Hence, in investigating the properties of gases according to this theory, we have to adopt what is called the statistical method. In this method we do not consider the behaviour of one particular molecule, but we take such a large number of molecules into consideration that, although the velocities of individual molecules may vary considerably, the *mean* velocity of all the molecules considered at any moment will be the same as the *mean* velocity of the same molecules, say one second later, or will be the same as the *mean* velocity of an equal number of other molecules of the same gas taken under the same conditions of temperature, pressure, &c. As an illustration of such a method, suppose cloth had to be bought to clothe an army of a million men, then, although the clothes made would be of many sizes, it is certain that the quantity of cloth used from year to year for this purpose would be the same, and we could calculate what is the quantity of cloth required to clothe the average-sized soldier. Instead, therefore, of attempting to allow for the various velocities of the different molecules, we shall suppose that they all move with the *mean* of their actual velocities. In the same way the length of the path traversed by each molecule between successive collisions varies greatly from time to time, but under given conditions the mean length of the path between successive collisions, or the mean free path, as it is called, will for any large number of molecules be the same, under similar conditions.

142*. Pressure Exerted by a Gas.—Suppose that a molecule of mass m moving with a speed V impinges at right angles on a solid surface, then it will, if it is perfectly elastic, rebound with a speed V, but in the opposite direction. The change in momentum of the molecule due to the impact will therefore be $2mV$. Hence, by § 60, the impulse of the blow on the surface is $2mV$. Suppose now we have a certain mass of a gas enclosed in an envelope, which for simplicity we may take to be a cube having each edge one centimetre long. The molecules in this vessel will be moving in all directions, but we may resolve the velocity of each molecule along three directions parallel to the mutually perpendicular edges of the cube; or, what comes to the same thing, if the number of molecules is very great, we may suppose that one-third of the total number of molecules are moving parallel to each of these three edges with the mean velocity V. Under these circumstances, the molecules of each group are moving parallel to four faces of the cube, and therefore will not impinge on them: they will only impinge on the two faces which are at right angles to their direction of motion. If we consider one molecule of one of these groups moving with the velocity $\bar{V}$, then the interval between two consecutive impacts of this molecule on *one* of the

faces will be the time taken by this molecule to travel to the opposite face of the cube and back again, that is, through a distance of two centimetres. Hence the interval between two consecutive impacts on the face will be $2/\bar{V}$, and there will be $\bar{V}/2$ impacts on the face by this molecule in each second. The impulse acting on the face due to each impact being $2m\bar{V}$, the total impulse during a second will be $m\bar{V}^2$, which is what would be produced by the action of a continuous force $m\bar{V}^2$, since the impulse of this force, if it acted for one second, would be $m\bar{V}^2 \times 1$. If the total number of molecules per cubic centimetre at the given pressure, &c., is N, then since $N/3$ molecules may be considered as moving parallel to the one considered above, the total force acting on the face will be $\frac{1}{3}Nm\bar{V}^2$. Since this force acts on unit area, if p is the pressure which the gas exerts on the containing wall, then

$$p = \tfrac{1}{3}Nm\bar{V}^2.$$

Now, since there are by supposition N molecules in the cubic centimetre, and the mass of each molecule is m, the total mass of the gas is mN, but the mass of unit volume of a body is the density ; hence, if ρ is the density of the gas,

$$p = \tfrac{1}{3}\rho\bar{V}^2,$$

or

$$p/\rho = \bar{V}^2/3.$$

Now according to Boyle's law pv = constant, if v is the volume of a given mass of gas. But the density of the gas is equal to M/v, or $v = M/\rho$, so that for a given mass of gas the volume is inversely proportional to the density, and Boyle's law may be written

$$p/\rho = \text{constant}.$$

Hence we see that, if Boyle's law holds, the mean velocity of the molecules $\bar{V}$ is constant.

From the equation $\bar{V} = \sqrt{\frac{3p}{\rho}}$, the value of $\bar{V}$ can be calculated if we know the density of a gas at any given pressure. Since the value of $\bar{V}$ is inversely proportional to the square root of the density, this result enables us to see why it is that the rate of diffusion of a gas is inversely as the square root of the density.

In the following table the values of $\bar{V}$, at a temperature of 0° C., are given for some gases :—

Hydrogen	. 185900 cm. per sec.
Nitrogen	. 49200 „ „
Oxygen	. 46500 „ „
Carbon dioxide	. 39600 „ „

143*. Avogadro's Law.—If we have two gases under the same pressure, and at the same temperature, N_1 being the number of mole-

cules per unit volume of one gas, m_1 the mass of each molecule, and V_1 the mean velocity of translation of the molecule; N_2, m_2, and $\bar{V}_2$ being the corresponding quantities for the other gas. Then, since the pressure is the same in each gas, we have, from the result obtained in the previous section,

$$\tfrac{1}{3}N_1 m_1 \bar{V}_1^2 = \tfrac{1}{3}N_2 m_2 V_2^2 \quad . \quad . \quad . \quad (1)$$

Now $\frac{1}{2}m_1\bar{V}_1^2$ is the kinetic energy of one of the molecules of the first gas when it is moving with the mean velocity. The mean kinetic energy of the molecules depends on the temperature, as we shall see later on. Also, if two gases are at the same temperature, the mean value of the molecular kinetic energy must be the same, for otherwise, when they are mixed, since now by collisions between the molecules the kinetic energy would become equalised, the temperature would alter. Thus the mean kinetic energy being the same for the gases, if the temperature is the same, we have

$$\tfrac{1}{2}m_1\bar{V}_1^2 = \tfrac{1}{2}m_2\bar{V}_2^2.$$

Combining this equation with equation (1) above, we get

$$N_1 = N_2,$$

or, under the same condition of pressure and temperature, equal volumes of all gases contain an equal number of molecules. This law was enunciated by Avogadro, who was led to it by purely chemical considerations.

The effect of temperature on the movements of the molecules of a gas will be considered in the section on Heat. Space and the scope of this book will not allow of our pursuing the subject of the kinetic theory of gases any further, and we must refer readers for further information on the subject to Clerk Maxwell's "Theory of Heat," or Risteen's "Molecules and Molecular Theory."

CHAPTER XVII

PROPERTIES OF LIQUIDS

144. Equilibrium of a Liquid at Rest.—In the case of a liquid at rest under the influence of gravity the free surface must be horizontal. If it were inclined to the horizon, then the weight of a particle P (Fig. 112) of the liquid at the surface would have a component parallel to the surface of the liquid. Since the surface is everywhere at the same pressure, there would be nothing in the nature of a hydrostatic pressure to resist this force, and as the liquid itself would offer no resistance, the particle P would move along the surface, and hence the liquid would not be at rest.

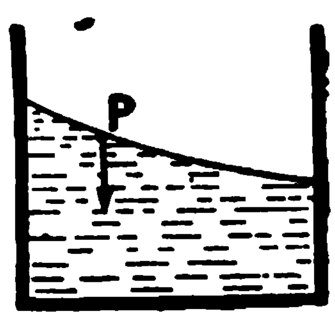


FIG. 112

Although a comparatively small surface of a liquid is for all practical purposes plane, it is not absolutely so, and when dealing with large surfaces, this departure from planeness becomes appreciable. The condition that the particle P (Fig. 112) should be at rest is that the line of action of the attraction of gravity on P should be normal to the surface at P. Hence the surface of a liquid is always normal to the radius of the earth at the point considered, and therefore forms part of a sphere having the earth's radius as radius. The fact that the surface of a liquid is always horizontal is made use of in the spirit-level. This consists essentially of a glass tube ABC (Fig. 113), which is slightly bent, and fitted, with the convex surface upwards, in a frame DE. This tube is closed at either end, and is filled with alcohol[1] except for a bubble of air B, which is left in. This bubble constitutes the only free surface of the liquid, and it always sets itself at the highest point of the curved tube. Hence, if the tube is fixed in the frame in such a way that when the lower surface of the frame is horizontal the highest point of the bent tube is opposite a fixed mark on the top of the tube, then, whenever the bubble is opposite this mark, the lower surface of the stand will be horizontal. If one end, say E, is tilted up, then the marked point of the tube is no longer the highest, and the bubble moves towards E.

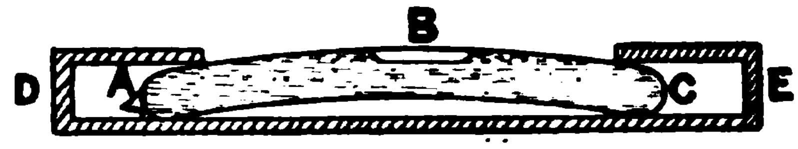


FIG. 113.

[1] Alcohol is used on account of its extreme mobility.

145. Level of Liquid Surface in Communicating Vessels.—Suppose a **U**-tube ABC (Fig. 114) has the same liquid in either limb, then the two surfaces A and B will be in the same horizontal plane. For if we consider a point D within the liquid, at a depth h_1 below the surface at A, and at a depth h_2 below the surface at C, then the pressure at D must be the same, whether caused by the column $\overline{AD}$ or the column $\overline{CD}$; otherwise the liquid would move towards the side on which the pressure was least. Hence

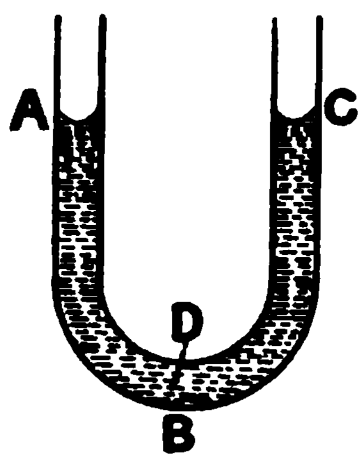


FIG. 114.

$$h_1 g\rho = h_2 g\rho,$$

where ρ is the density of the liquid.

$$\therefore \qquad h_1 = h_2.$$

By an exactly similar line of argument it can be shown that the pressure at any pair of points, one in either limb, is the same if these points lie in the same horizontal plane.

If, instead of having the same liquid in both limbs, one limb AB (Fig. 115) contains a liquid of less density than that in the other; then, if B is the surface separating the two liquids, from what has been said in the previous paragraph, the pressure at a point D in the denser liquid in the same horizontal plane as B must be equal to the pressure at B. Hence the pressure exerted by the column $\overline{AB}$ of the one liquid must be equal to the pressure exerted by the column CD of the other liquid. So that, if h_1 and h_2 are the heights of these columns, and ρ_1 and ρ_2 are the densities of the liquids, we have—

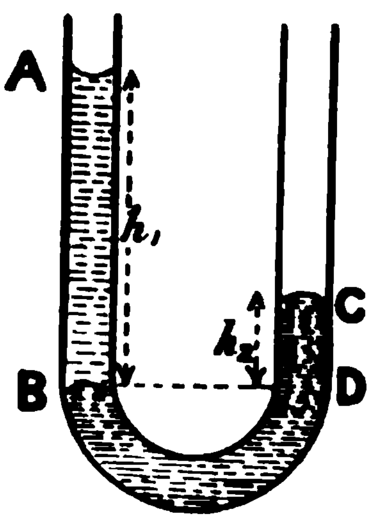


FIG. 115.

$$h_1\rho_1 g = h_2\rho_2 g,$$

or

$$\frac{h_1}{h_2} = \frac{\rho_2}{\rho_1}.$$

That is, the heights of the columns of the two liquids above the level of their common surface are to one another inversely as the densities of the liquids.

146. Density of Liquids.—In order to determine the density of a liquid, the mass of a known volume must be measured. If, however, we know the density of any given liquid, say water, then we can find the density of any other liquid by comparing the mass of any volume of the liquid with that of the same volume of the standard liquid.

The density of water has been determined by Macé de Lépinay with great accuracy, by a method depending on the principle of Archimedes. A cube of quartz was prepared and its volume obtained by measurement.

The planeness of the faces was tested, and the distance between the opposite faces measured by an optical method depending on the production of Newton's rings (see § 376). This cube was then placed on the pan of a very delicate balance, a small cage suspended by a fine platinum wire hanging from the under side of the same pan. This cage was completely immersed in the water of which the density was to be measured, and which was kept at a constant temperature, this temperature being read by means of a thermometer. The weights necessary to counterpoise the quartz block (in air) and the wire cage (immersed in the water) having been placed in the second pan of the balance, the quartz block was transferred to the cage, and the weights found which were now necessary to counterpoise. The difference between these weights represents the loss of weight of the block when immersed in water, and this, by the principle of Archimedes, is equal to the weight of a volume of water equal to that of the block. Hence, knowing the volume of the block, *i.e.* the volume of the water displaced, the density can be calculated. The object of having the wire cage, which was always immersed in the water, was to allow for the weight of the suspending fibre and that of the water it displaced; also, by this arrangement the effect of the surface of the liquid on the suspending wire due to capillarity (§ 157) was the same during both the weighings, and was therefore eliminated.

Since the volume of the quartz cube altered with the temperature, this had to be allowed for, so that a preliminary measurement of the coefficient of expansion of quartz was made.

The following table gives the density of water at different temperatures :—

DENSITY OF WATER.

Temperature.	Density.	Temperature.	Density.
Deg. C.	Grams per c.c.	Deg. C.	Grams per c.c.
0	.99984	16	.99894
1	89	17	77
2	93	18	59
3	95	19	40
4	96	20	19
5	95	21	.99798
6	93	22	76
7	89	23	52
8	84	24	28
9	77	25	03
10	.99969	26	.99677
11	60	27	50
12	49	28	22
13	37	29	.99593
14	24	30	63
15	09		

Knowing the density of water, Δ, at different temperatures, we can determine the volume of a solid which is insoluble in water, by weighing it first in air and then when immersed in water at a known temperature. If w_1 is the weight in air and w_2 the weight in water, then the loss of weight, that is, the weight of water displaced, is $w_1 - w_2$, and this must be equal to ΔV, where V is the volume of the solid. Thus

$$V = (w_1 - w_2)/\Delta.$$

One method of comparing the density of a liquid with that of water is to determine the loss of weight of a solid, which is unacted upon by either liquid, when weighed first in water and then in the liquid. In this way the weights or masses of equal volumes of the liquid and of water are obtained. If m_1 is the loss of weight in the given liquid of density ρ, and m_2 is the loss of weight in water of density Δ, then

$$m_1 = V\rho$$

and

$$m_2 = V\Delta,$$

where V is the volume of the solid. Hence

$$\rho = \frac{m_1}{m_2}\Delta;$$

and, by taking the value of Δ for the temperature of the experiment from the table given above, ρ can be calculated.

Another method in common use for determining the density of a liquid is to weigh a small bottle, called a specific gravity bottle or pyknometer, when full, first of water, then of the liquid. Two forms of pyknometer which are in common use are shown in Fig. 116. One consists of a small glass bottle A, fitted with a well-ground-in glass stopper B. This stopper is pierced by a fine hole. The bottle is completely filled with the liquid, and the stopper inserted, care being taken not to include any air-bubbles. The superfluous liquid flows out through the hole in the stopper and is wiped off. The other form consists of a bent glass tube CDE of the shape shown. The end C is drawn out into a fine capillary, and a fine mark is engraved on the glass at F. To fill the tube the end C is dipped below the surface of the liquid, which is drawn into the tube by suction at E till it fills it a little above the mark F. Then, by touching the capillary C with a piece of blotting-paper, liquid is withdrawn till the surface comes exactly to the mark F.

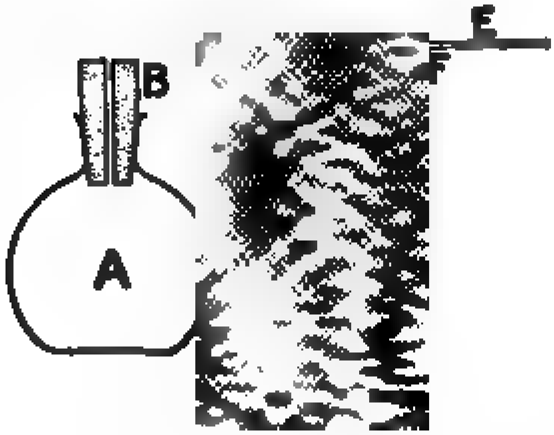


FIG. 116.

Let w_1 be the weight of the empty pyknometer, and w_2 and w_3 the weight when full of the liquid and water respectively. Then $w_3 - w_1$ is the weight of water which fills the instrument. Hence if Δ is the density of the water at the temperature at which the pyknometer was filled, its volume V is given by

$$V = (w_3 - w_1)/\Delta.$$

The weight of a volume V of the given liquid is $w_2 - w_1$. Hence the density ρ of the liquid is

$$\rho = \frac{w_2 - w_1}{w_3 - w_1} \cdot \Delta.$$

The following table gives the density of some liquids :—

DENSITY OF LIQUIDS.

Liquid.	Mass of 1 c.c.	Temperature.
	Grams.	Degree C.
Mercury	13.596	0
Bromine	3.187	0
Chloroform	1.480	18
Glycerine	1.260	0
Milk (cow's)	1.03 (about)	15
Sea water	1.025	15
Olive oil	0.918	15
Turpentine	0.873	16
Alcohol (ethyl) . . .	0.791	0
Ether	0.736	0

147. Flotation.—Since when a body is immersed in a fluid it experiences an upward force, due to the pressure of the fluid, equal to the weight of the fluid displaced, it follows that if the density of the body is less than that of the fluid, the weight of the displaced fluid will be greater than the weight of the body, and hence the upward force acting on the body due to the fluid will be greater than the downward force due to gravity. If no other forces are acting on the body, it will therefore rise until the weight of the displaced fluid is exactly equal to that of the body. In the case of a body such as a balloon in the air, this will happen when it has risen to such a distance that the density of the air has become so much reduced that the weight displaced by the balloon is equal to its own weight. In the case of a solid immersed in a liquid, it will rise till, some of the solid having risen above the surface of the liquid, the weight of the volume of liquid displaced by the remainder, which is still submerged, is equal to the weight of the body.

In order that a body floating in a liquid may be in equilibrium, not only must the upward pressure due to the liquid be equal in magnitude to the weight of the body, but it must also act vertically upwards through the centre of gravity of the body. If we suppose the body removed and

M

replaced by some of the liquid which has become solid and occupies exactly the position of the *immersed* part of the solid, this solidified portion of liquid will be in equilibrium. Hence, since its weight acts vertically through its centre of gravity, the pressure due to the part of the liquid which has remained liquid must also act vertically through the centre of gravity of the solidified portion. The direction and magnitude of this pressure must be the same as that which was acting on the floating body, so that we see that the upward force due to the liquid is equal to the weight of the liquid displaced, and acts in a vertical direction through the point which would be the centre of gravity of the *displaced liquid.* If G (Fig. 117, A) is the centre of gravity of a floating body, and H the

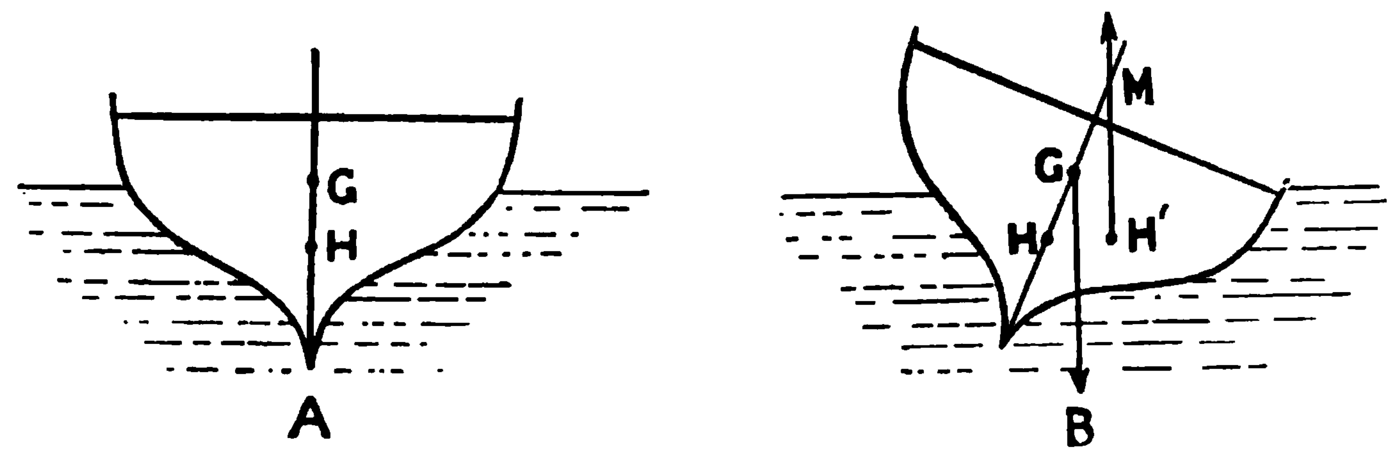


FIG. 117.

centre of gravity of the displaced water, the two points G and H must, if the body is in equilibrium, be vertically one over the other. If the body be displaced into some such position as that shown at B, then the centre of gravity of the displaced liquid will no longer be at H, but at some point such as H′. The body is then acted upon by a couple which tends to bring it back into the position shown at A. The point M, where the new vertical through H′ cuts the vertical drawn through H in the undisturbed position, is called the *metacentre.* In the above figure the metacentre is above the centre of gravity, and the floating body is in stable equilibrium, as the couple when it is deflected tends to restore it to its original position. In the case shown in Fig. 118, however, the couple, which comes into play when the body is deflected from the A position, tends to increase the displacement, and hence the condition figured at A is unstable. In this case it will be seen that the metacentre M is *below* the centre of gravity of the floating body. Hence a floating body is stable when the metacentre is above the centre of gravity, and the higher the metacentre is above the centre of gravity, the more stable is the body. If the metacentre coincides with the centre of gravity of the floating body, as it does in the case of a sphere, the body is in neutral equilibrium, while

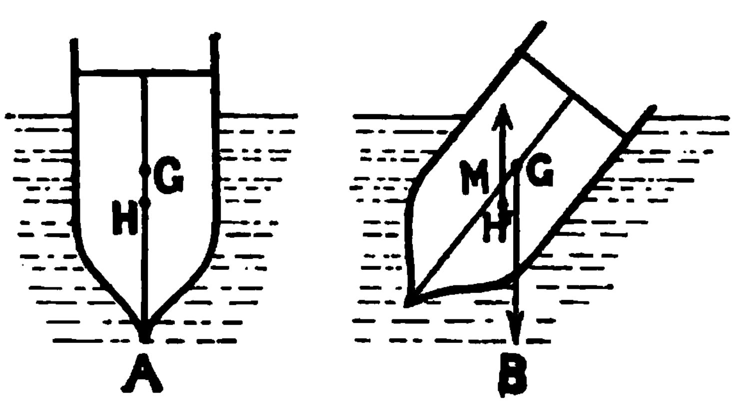


FIG. 118.

if the metacentre is below the centre of gravity the equilibrium is unstable. These principles are of great importance in designing ships, the object of ballast being to lower the centre of gravity so as to keep it well below the metacentre.

148. Hydrometers.—The volume of a floating body immersed in a liquid depends on the density of the liquid, for the product of the density of the liquid into the volume of the body immersed, which gives the weight of liquid displaced, must always be equal to the weight of the body. Hence the volume of a body immersed in different liquids may be employed to compare the densities of liquids. An instrument depending on this principle, called a hydrometer, is shown in Fig. 119. It consists of a glass bulb B, to the lower end of which a small bulb A is fixed, and at the upper end a narrow glass stem CD. Some mercury or lead-shot is placed in A, so that the instrument floats upright. The stem CD is graduated, so that the depth to which the instrument sinks in the different liquids can be read off. Suppose that the division to which it sinks, when floating in water, is marked 100, and the volume of the stem included between two consecutive divisions is 1/100th of the total volume immersed when the body floats in water. Then, if the instrument when floating on a liquid sinks to the 60th division, this means that 60 units of volume of this liquid have the same weight as 100 units of volume of water. Hence the specific gravity of the liquid is 100/60 or 1.67. In the same way, if the instrument sinks in a liquid to the 120th division, the specific gravity is 100/120 or 0.833. The length of the stem, which will have a volume equal to 1/100th of the volume of the instrument, varies inversely as the cross section, so that by making the stem very narrow the sensitiveness can be increased. We cannot, however, increase the sensitiveness indefinitely, on account of the force which the surface of the liquid exerts on the stem (see § 157), and which prevents the instrument taking up the position it would if no such capillary force existed. In order to avoid having a very long stem, it is usual to have a series of hydrometers, so weighted that in the liquid in which one sinks to nearly the top of its stem, the next only sinks to a point near the bottom of its stem. The stems of hydrometers are often graduated so as to give the specific gravity of the liquid in which they are placed directly. In this case the length of the graduations is not constant, but decreases from the top to the bottom of the stem. The position of two points on the scale are generally found by floating the instrument in two liquids of known specific gravity, and from the position of any two such points on the scale the rest of the graduations can be obtained.

FIG. 119.
(From Watson's *Elementary Practical Physics.*)

Another form of hydrometer is shown in Fig. 120, and is called Nicholson's hydrometer. In this instrument the stem is not graduated, but has a single mark, O, and when in use the instrument is always sunk to this mark, so that the volume immersed is constant. Attached to the float A are two scale-pans B and C, the lower one being weighted so that the instrument can float upright.

When using this instrument to determine the density of a liquid, it is first floated in water at a known temperature, and weights are placed in the upper scale-pan till the mark O is coincident with the surface of the liquid. Let W be the weight of the instrument itself, and w_1 the weights added; then the weight of the water displaced is $W+w_1$, and the volume V of the displaced water is given by

FIG. 120.

$$V=(W+w_1)/\Delta,$$

where Δ is the density of the water at the given temperature.

Now let the instrument be floated in a liquid of density ρ, and let the weight which has to be placed in the pan B to bring the mark to the surface of the liquid be w_2. Then the weight of the liquid displaced is $W+w_2$. Now, since the volume of the hydrometer immersed is the same as before, namely V, we have—

$$\rho=(W+w_2)/V.$$
$$=\frac{W+w_2}{W+w_1}\cdot\Delta.$$

This instrument is more often used for finding the density of solids than of liquids. For this purpose the hydrometer is floated in water, and the solid placed in the pan B, and weights w_3 added till it sinks into the sighted position. Since, when the solid is not present, the weight necessary to sink the instrument is w_1, it follows that the weight of the solid is w_1-w_3. Next, the solid is placed in the lower pan C, and the weight w_4 necessary to sink the instrument determined. The solid being immersed in water, will lose in weight an amount equal to the product $V\Delta$, where V is its volume. Hence

$$V\Delta=w_4-w_3,$$

or

$$V=(w_4-w_3)/\Delta.$$

Therefore the density ρ of the solid is given by

$$\rho=\frac{w_1-w_3}{w_4-w_3}\Delta.$$

149. Elasticity of Liquids.—Liquids can only have bulk elasticity, *i.e.* they only offer resistance to change of volume. As has already been mentioned, liquids require a very great pressure to reduce their bulk

appreciably, differing in this respect very markedly from the other division of fluids, namely, gases. In fact, a liquid is sometimes defined as a fluid which offers great resistance to change of volume. The first person to show that liquids were compressible was Canton in 1761, while the first measurements of any accuracy were made by Œrsted (1822).

The instrument used by Œrsted, and called a piezometer, is illustrated in Fig. 121. It consists of a strong glass cylinder, to the top of which is cemented a metal cap. Water can be introduced into the cylinder through the funnel R, and by turning the screw P a small piston is forced down, thus compressing the water in the cylinder. Within the cylinder is a glass vessel A filled with the liquid to be experimented on, and terminated by a capillary tube, the open end of which dips beneath the surface of some mercury, O. The volume of the vessel A, as well as the volume of unit length of the capillary, having been determined by filling with mercury and weighing, the total volume of the liquid in A, and the decrease in volume corresponding to any observed rise of the mercury into the capillary tube, when pressure is applied, is known. The pressure is measured by means of an air manometer (§ 131) B, consisting of a glass tube closed at the top, with its open end below the surface of the mercury. This tube, before the pressure is applied, is quite full of air. When the pressure is applied the mercury rises, the air being compressed, so that by noting the change in volume by means of the attached scale, the pressure can be calculated.

FIG. 121.
(From Ganot's "Elementary Treatise on Physics.")

Œrsted assumed, since the glass vessel A is subjected to the same pressure both inside and out, that therefore its volume was the same at all pressures, and hence that the contraction observed was due solely to the compression of the *liquid*. This assumption is, however, not justifiable. If we imagine the vessel to be in the form of a sphere, and that the pressure applied both inside and out is, say, 100 atmospheres. Then suppose that the same pressure were applied externally to a *solid* glass sphere of exactly the same external dimensions as the hollow sphere. The solid sphere would contract, since, as will be seen later, glass is compressible; and if we consider a thin external layer of the sphere, it will eventually be at rest under the pressure of 100 atmospheres acting everywhere towards the inside, and an equal and

opposite pressure due to the elasticity of the inside layers of the solid sphere. The condition of this layer, therefore, resembles that of the hollow glass sphere filled with liquid and subjected inside and out to a pressure of 100 atmospheres, and since it contracts in volume under these conditions, we infer that the hollow sphere would do so also.

In order to overcome this difficulty, Regnault devised a modified form of piezometer, in which the change in volume of the containing vessel could be observed. His apparatus consisted of a strong glass bulb A (Fig. 122) with a fine capillary neck B. This bulb, in which the liquid to be tested is contained, was inside a strong metal vessel C. The tube D was connected with a compression-pump, and by means of the taps E, F, G, H the volume of the liquid in the bulb was observed under the following conditions: (1) The taps E and H being closed and F and G open, the apparent increase of the volume of the liquid due to the action of an external pressure p on the outside of the bulb was obtained; (2) the taps F and H being open and E and G closed, the apparent contraction of the liquid under the influence of a pressure p acting both on the inside and outside of the bulb, as in Œrsted's experiment, was obtained; (3) the taps E and H being open and F and G closed, the apparent contraction of the liquid when the pressure p acted on the inside of the vessel only was obtained. From the results of these three observations, the effect of the contraction of the envelope can be eliminated and the true coefficient of compressibility, *i.e.* the diminution produced in unit volume by unit increase in pressure, can be calculated, although even in this case some assumptions as to the uniformity of the thickness of the walls of the vessel have to be made. In some of his experiments Regnault, therefore, used spherical bulbs of brass or copper.

FIG. 122.

The following table (calculated by Tate from Cailletet's results) gives the coefficient of compressibility per atmosphere for some liquids :—

Liquid.	Temperature.	Pressure in Atmospheres.	Coefficiency of Compressibility per Atmosphere.
Water . . .	8° C.	705	0.0000469
Sulphuric ether	10° C.	630	0.0001458
Bisulphide of carbon	8° C.	607	0.0000998
Sulphurous acid .	14° C.	606	0.0003032

It will be seen that, as compared with gases in which the coefficient of compressibility at a pressure of one atmosphere is 0.5, liquids are very little compressible.

150. Hydraulic Press.—Pascal's law, that liquids transmit in all directions and without diminution any pressure that is applied to them, receives an important application in the hydraulic press or ram. This machine was invented in 1795 by Bramah, and is shown in sections in Fig. 123. It consists of a large metal cylinder, A, with very strong sides, in which an iron ram works water-tight, through a joint, B. This joint is made water-tight by means of a circular leather washer, the section of which is U-shaped, the concave surface being turned towards the inside of the cylinder. The pressure of the water in the cylinder forces this washer against the ram

FIG. 123.

on the inside, and against the neck of the cylinder on the outside, so that the greater the pressure of the water the more tightly does the washer fit. The cylinder is connected by a strong pipe, C, with a force-pump, of which the piston, D, is of small diameter. By this means water can be pumped into the large cylinder A. When the plunger of the pump is forced down, the liquid in the machine transmits the pressure to the base of the ram, which is forced up. If a is the area of cross section of the plunger of the pump, and the downward force exerted on the plunger is P, then the pressure exerted on the water in the pump is P/a. This pressure is transmitted to the cylinder A, and hence a pressure of P/a acts on each unit of surface of the base of the ram. If A is the area of cross section of the ram, the total upward force exerted on it is AP/a. In other words, the force (W) exerted by the ram is to the force acting on the plunger of the pump as the area of cross section of the ram is to that of the plunger, or

$$\frac{W}{P}=\frac{A}{a}=\frac{D^2}{d^2}$$

if D is the diameter of the ram and d the diameter of the plunger.

The principle of the hydraulic press is also employed as a means of storing power, which is required in an intermittent manner, such as for working lifts. In this case powerful pumps are employed to pump water into a strong steel reservoir fitted with a wide piston, like the ram of a Bramah press, which is loaded with heavy weights. The work

performed by the engine which drives the pumps is employed in raising these weights, and the potential energy thus stored up can be usefully employed by connecting the reservoir by pipes to the hydraulic engines to be driven by the water under high pressure.

151. Pumps.—Pumps may be divided into two classes—suction-pumps and force-pumps—though most pumps which are used in practice really consist of both kinds combined, the suction or the force, as the case may be, however, generally playing quite a subordinate part.

The air-pump described in § 136 (Fig. 105) is a suction-pump, and if the tube leading to the receiver were connected to a tube the end of which dipped beneath the surface of a liquid, when the piston was raised the pressure of the air acting on the surface of the liquid would force it up past the valve B into the cylinder. When the piston descends, the valve B shuts and C opens, the liquid escaping on to the top of the piston. If a spout were fixed to the top of the barrel of the pump the liquid would flow out through E, and we should have the ordinary domestic pump. Since it is the pressure of the atmosphere which drives the liquid up into the barrel of the pump when the piston rises, it follows that a suction-pump cannot raise a liquid through a greater height than that of the column in a barometer filled with the liquid. Hence if ρ is the density of the liquid, and h is the height in centimetres through which the liquid is to be raised, a suction-pump cannot be used if ρh is greater than 13.6×76, since 13.6 is the density of mercury, and 76 cm. is the height of the mercury barometer. If the liquid is water, $\rho = 1$, and the limiting value for h is 13.6×76, or 1033.6 cm., or about 34 feet.

FIG. 124.

In the force-pump, the second valve, instead of being placed in the piston P (Fig. 124), is at D. When the piston rises, the liquid enters the barrel of the pump through the tube B and the valve C, the valve D remaining shut. When the piston descends the valve C shuts, and the liquid is forced out through the valve D and up through the tube E. The height to which the liquid can be forced by this form of pump is only limited by the strength of the barrel A and the force that can be applied to drive the piston down. The pumps illustrated in Figs. 99 and 123 are force-pumps. In the fire-engine, which is a force-pump, there is an air-chamber A (Fig. 125), the water being forced in through B by the pump during the down-stroke of the piston, and escaping through C. The pipe C is made so small that the water cannot escape as fast as it is pumped in during the down-stroke of the piston, so that the air which is enclosed in A becomes compressed, and so during the up-stroke of the piston this air, by expanding, continues

FIG. 125.

to drive the water out through C ; and in this way keeps up a continuous stream.

152. The Syphon.—The syphon consists of a bent tube ABC (Fig. 126), open at both ends, one leg being of greater length than the other. If the tube is filled with a liquid, and the end of the shorter limb dipped beneath the surface of some of the liquid, then the pressure at the end A, tending to force the liquid into the tube, is equal to the atmospheric pressure minus the weight of the column of liquid $\overline{DB}$. The pressure at C, tending to force the liquid up the tube, is the atmospheric pressure less the weight of the column of liquid $\overline{CB}$. Hence, since $\overline{CB}$ is greater than $\overline{DB}$, the pressure tending to force the liquid in at A is greater than that at C, so that the liquid will flow in at A and out at C as long as the surface of the liquid is above the end A. The syphon depending, as it does, on the atmospheric pressure to force the liquid up from D to B, will not work in a vacuum, or if the height $\overline{DB}$ is greater than that of the barometric column of the given liquid. In either of these cases the liquid would not fill the bend of the syphon, but there would be a vacuum there.

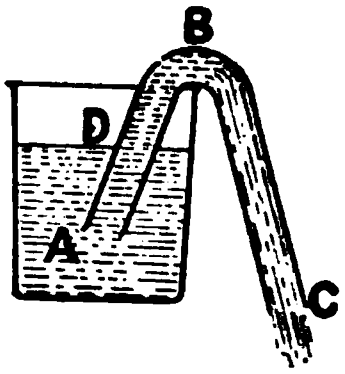


FIG. 126.

153*. Kinetics of Liquids—Law of Continuity.—In the case of a liquid flowing in a pipe, the volume of liquid that passes across any cross section of the pipe during any given time must be the same. If this were not so, then either the liquid would accumulate or diminish between the two sections. Thus if we have a liquid flowing through a pipe of variable cross section (Fig. 127), the volume of liquid which crosses each of the sections A, B, C, or D, during a given interval must be the same.

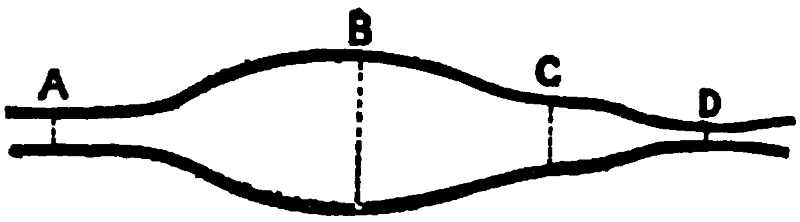


FIG. 127.

It follows from this that the velocity of the liquid is everywhere inversely proportional to the area of cross section. For if a is the cross section of the tube at A, and v_a is the velocity with which the liquid is moving, the volume which passes A in unit time is av_a. In the same way the volume which passes B in unit time is bv_b, where b is the area of cross section at B, and v_b is the velocity of the liquid. Since the quantity of liquid which crosses A in unit time must be equal to the quantity which crosses B, we get

$$av_a = bv_b,$$

or

$$\frac{v_a}{v_b} = \frac{b}{a}.$$

154*. Force producing Motion in a Liquid.—In every case a liquid tends to flow from a point at which the pressure is high to a point a

which it is lower. Hence if we find that the velocity of flow of a liquid increases as we move from one point to another, it shows that the pressure at the first point is greater than at the second. For the fact that the velocity of the liquid is increasing, that is, has a positive acceleration, shows that some force must be acting on it tending to increase its velocity, and the only force acting is the difference in the pressures at the points considered. In the case of the liquid flowing through the tube shown in Fig. 127, the velocity at A is greater than the velocity at B. Hence the velocity of the liquid diminishes from A to B, and therefore the liquid must be moving against a force, or the pressure at B must be greater than at A. In the same way the pressure at D is less than at A, B, or C.

If we consider a liquid in which the particles exert no friction one against another, flowing in a tube without friction against the walls, then the force producing motion on a small cube of the liquid of which the edges are parallel to the direction of flow of the liquid, is the difference of pressure between the ends of the cube at right angles to the direction of motion. If p_1 and p_2 are these pressures, and s is the length of the edge of the cube, the force causing motion is $(p_1-p_2)s^2$, for s^2 is the area of a face of the cube, and if ρ is the density of the liquid, the mass of the cube of liquid is $s^3\rho$. Hence the acceleration (a) with which the liquid moves is given by the equation—

$$(p_1-p_2)s^2=s^3\rho\,.\,a,$$

or

$$a=\frac{p_1-p_2}{s}\cdot\frac{1}{\rho}.$$

The quotient $(p_1-p_2)/s$ is the rate at which the pressure decreases with distance measured parallel to the direction of flow, and is called the *pressure gradient* or *pressure slope* in this direction. The flow of a liquid always takes place in the direction in which the pressure gradient is greatest, and the greater the pressure gradient, the greater the velocity of the liquid.

When a liquid flows out of a vessel through an opening, the pressure just outside the opening is zero, while the pressure just inside is that due to a column of liquid of the height of the free surface of the liquid in the vessel above the opening. This height of the free surface above the opening is called the *head* of the liquid which produces the pressure, and the pressure is equal to $H\rho g$, where H is the head and ρ is the density of the liquid.

155*. Velocity of Outflow of a Liquid (Torricelli's Law).—In order to obtain an expression for the velocity with which a liquid will flow out, under the action of gravity, through an opening in the base of a vessel containing the liquid, let us assume that the cross section of the opening is a, and that of the vessel at the free surface of the liquid is A, while

the head of liquid is H. Let v be the velocity[1] with which the liquid, of density ρ, escapes from the opening. Suppose that, in the very short time δt, the level of the free surface of the liquid falls through a distance h (Fig. 128), then the volume of liquid that has escaped in this time is Ah. Hence the potential energy of the liquid within the vessel has decreased by the amount $Ah\rho.Hg$, for each layer of the liquid has fallen through a height h, or, in other words, the whole volume of liquid HA has fallen through this distance. The kinetic energy of the liquid as it escapes is $\frac{1}{2}Ah\rho v^2$, and since the kinetic energy gained must be equal to the potential energy lost, we have :—

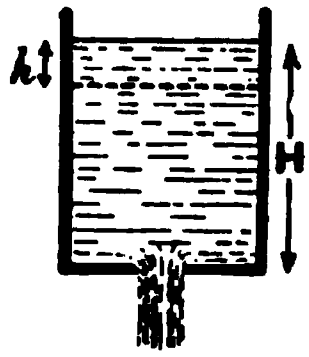


FIG. 128.

$$\tfrac{1}{2}Ah\rho v^2 = Ah\rho Hg.$$

$$\therefore \qquad v^2 = 2gH,$$

or $$= \sqrt{2gH}.$$

That is, the velocity with which the liquid escapes varies as the square root of the head, and it will be noticed that the velocity is independent of the density of the liquid and of the size of the opening, and is the same as that which would be acquired by a body falling freely through the height H, *i.e.* through the distance between the free surface of the liquid in the vessel and the opening.

If the opening, instead of being made through the bottom of the vessel, is made through the side, the stream of liquid will be projected horizontally with a velocity $\sqrt{2gH}$. Hence each particle of the liquid will move in a horizontal direction with a uniform velocity $\sqrt{2gH}$, and also with a uniform acceleration g in the vertical direction, so that the liquid jet will form a parabola (§ 40). If the opening is made so that the liquid jet is directed upwards, the highest point reached by the water ought, according to the above calculation, to be in the same horizontal plane as the free surface of the liquid in the vessel. On account, however, of viscosity (§ 162) and the resistance of the air, the liquid never quite rises to the level of the surface.

The volume of liquid which will escape per second through an opening of cross section a is found experimentally to be less than va. This is due to the fact that the jet of liquid, as it leaves the opening, becomes contracted, forming the *vena contracta* as at A (Fig. 129), so that the cross section of the jet is less than that of

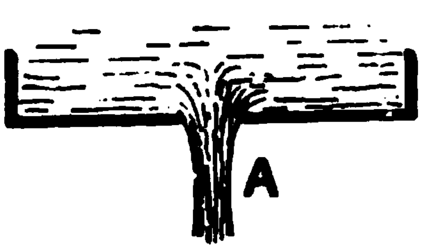


FIG. 129.

[1] By velocity of the liquid is meant the speed with which a small speck of dust would be carried along by the liquid.

the opening through which it passes. The actual volume of liquid that escapes is only about 62 per cent. of the volume calculated from the expression $a\sqrt{2gH}$, while the cross section of the *vena contracta* is about .62 times the cross section of the opening. The quantity of liquid which escapes can be considerably increased if a small cylindrical tube or ajutage, of the same diameter as the opening, is fitted to the aperture. In this case the outflow may be increased to about 82 per cent. of the calculated amount. If the ajutage is of considerable length, the outflow is again reduced, this being due to viscosity, *i.e.* friction between the different parts of the liquid.

CHAPTER XVIII

MOLECULAR PHENOMENA IN LIQUIDS

156. Cohesion.—If a rod or tube of glass is dipped into water and is then withdrawn, a drop of the liquid will be left hanging to the end of the rod. If more water is carefully added, the size of the drop will increase until its weight is sufficient to tear it away from the glass. In the same way, if a clean metal ring is dipped into a solution of soap and then withdrawn, a film of the liquid will remain stretched across the ring. In both these cases the effects are said to be due to the *cohesion* of the liquid. The term adhesion is, however, sometimes used to indicate the attraction manifested between a liquid and a solid, and the term cohesion restricted to the attraction between the different particles of a mass of liquid. This cohesive force is in most cases masked by the action of gravity, and hence to observe its effects we require to reduce the effects of gravity to a minimum.

Thus, if a large drop of oil is placed on the surface of water it immediately spreads. If, however, a mixture of alcohol and water is prepared of exactly the same density as the oil, and a drop, or even a considerable volume, of oil is introduced in the water, it immediately gathers itself into a sphere which remains suspended in the alcohol and water. By floating the oil in a liquid of the same density as itself we remove it from the influence of gravity, and then the cohesion between the liquid particles causes the drop to assume the spherical form.

157. Surface Tension.—In the case of the globule of oil floating in a liquid of the same density, the shape assumed is the same as the oil would take had it been enclosed in an elastic membrane or skin. The presence of such an elastic skin would also serve to explain the formation of the drops on the end of the glass rod or the soap film.

We can explain these facts on the molecular hypothesis in which it is assumed that in a liquid the molecules exert on one another an attractive force; this force, however, being only appreciable when the molecules are within a short distance of one another, which is called the range of molecular attraction. If we describe a sphere with any particle as centre having a radius equal to the range of molecular attraction, then we may neglect the effects of all the molecules which lie outside this sphere on the molecules at the centre.

In the case of a molecule A (Fig. 130) well within a liquid, the whole sphere will lie within the liquid, and hence the molecule A will be attracted by the neighbouring molecules equally in all directions. If, however, the molecule (B) is so near the surface of the liquid, E F, that the sphere would intersect the surface, then the attraction exerted on the molecule is not the same in all directions. The attraction due to molecules within that portion of the sphere in the liquid which is unshaded, being symmetrical about the molecule B, will have a resultant which is zero. The attractions of the molecules within the shaded part will, however, have a resultant directed towards the inside of the liquid mass, and perpendicular to the surface. In the case of a molecule actually on the surface, as at C, this resultant is a maximum. The effect of these unbalanced molecular forces acting on the molecules near the surface is to exert a pressure on the interior of a liquid mass, similar to that which would be caused by an elastic skin, and it is frequently convenient to speak as if such an elastic skin really existed, and to say that this pressure within a liquid mass is due to the *surface tension* of the liquid.

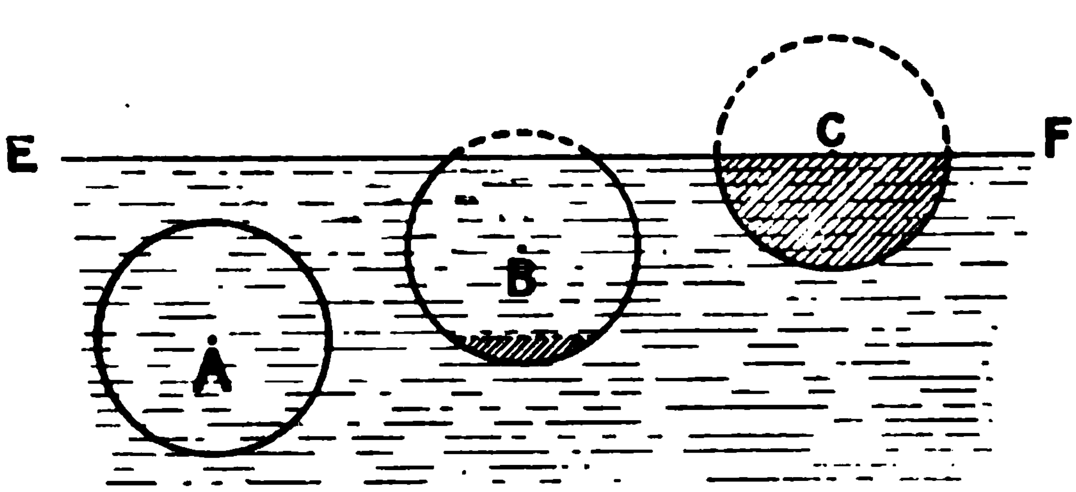


Fig. 130.

The magnitude of the pressure due to the surface tension depends on the form of the liquid surface. Let us take the case of three molecules,

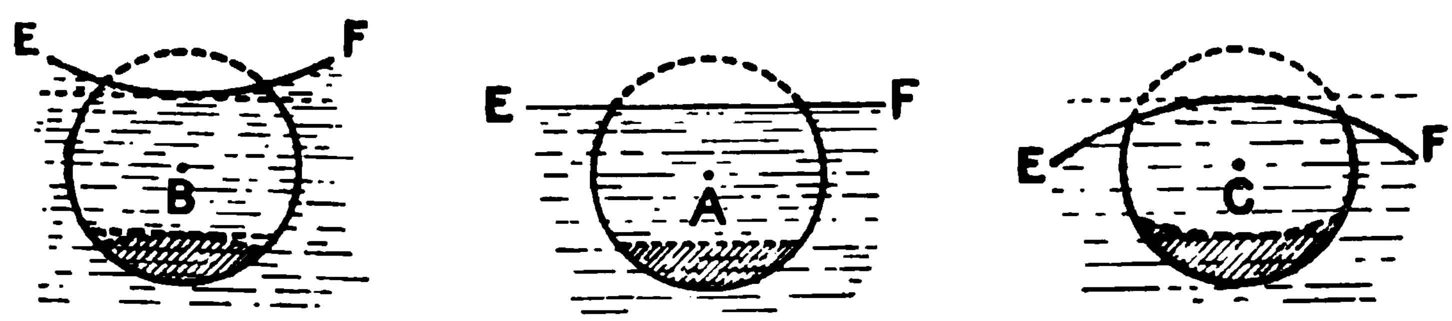


Fig. 131.

A, B, and C (Fig. 131), at equal distances, less than the radius of molecular attraction, from the surface EF, which in the first case is plane, in the second concave, and the third convex, and, as before, let us indicate by shading the part of the sphere of molecular attraction which is efficacious in producing an inwardly directed force on the molecule. If the surface is concave as at B, then, although the molecule B is at the same distance below the surface as is A, where the surface is plane, the shaded part is

less, so that the molecular force acting on B towards the inside of the liquid is less than that on A. In the case where the liquid surface is convex (C), the shaded part is larger than in A, and hence the force is larger. Looking at it from the point of view of an elastic membrane, it is evident that at B the elasticity of the membrane would diminish the pressure within the liquid, while at C it would increase the pressure.

The existence of this pressure due to molecular, as distinct from gravitational attractions, cannot be experimentally demonstrated, but there are many striking phenomena depending on the fact that the surface of a liquid is in a state of tension. Thus if a metal ring is dipped in a solution of soap, and a small loop of cotton, which has been previously moistened with the solution, is placed on the film left on the ring, this loop can be made to take up any form such as A (Fig. 132), and will retain this form. If, however, the film *within* the loop is broken, the loop immediately takes up the circular form shown at B; and if it is now deformed in any way, on being released it immediately springs back to the circular form. This behaviour is due to the fact that, in the first case, the surface tension of the liquid film acts equally on both sides of the cotton, but when the film inside the loop is broken, the surface tension only acts on one side, and hence draws the loop out into a circle. Another method of showing the surface tension is by means of a bent wire ABC (Fig. 133) and a straight wire DE, which simply rests against this. If a soap film is formed in the enclosed space DBE, it will be found that the surface tension acting on DE is able to support not only the weight of the wire DE, but also a small weight W. This arrangement might also be used to obtain a rough measure of the amount of the surface tension. If W is the mass of the cross wire DE and its attached weight, then the surface tension of the film supports weight W, and therefore exerts a force of Wg units of force. The surface tension of the film acts all along the portion of the wire DE, intercepted between the legs of the bent wire, and acts at right angles to the wire. Since the film has two surfaces, if the force exerted on unit length of DE due to the surface tension of *one* side of the film be T, then the whole upward force on

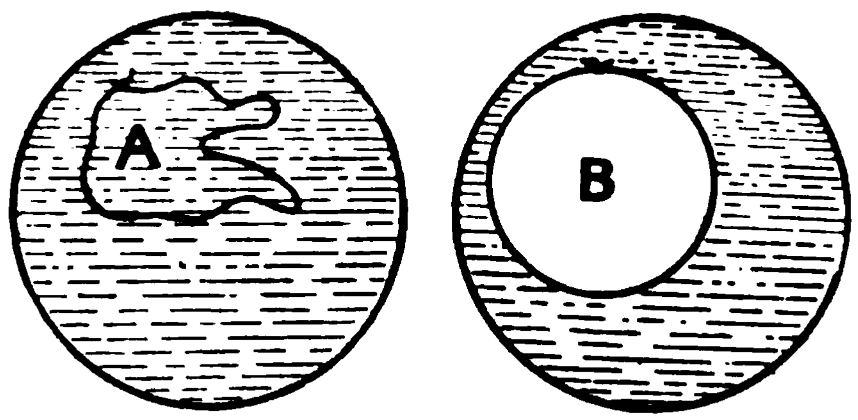


Fig. 132.

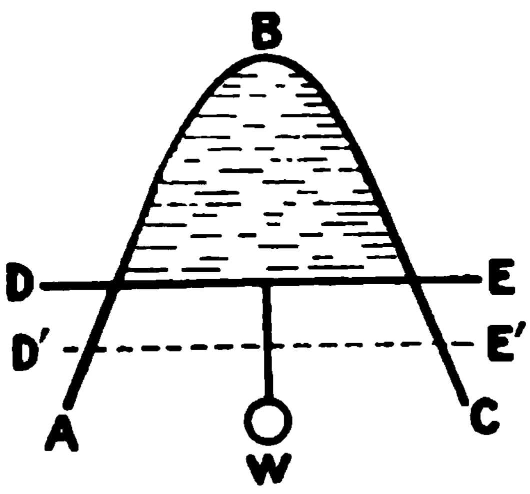


Fig. 133.

DE due to surface tension is $2Tl$, where l is the length of DE in contact with the film. Hence if there is equilibrium

$$2Tl = Wg,$$

or

$$T = \frac{Wg}{2l}.$$

The quantity T is called *the* surface tension of the liquid, and is the force exerted across unit length taken along the surface of the liquid. In the *c.g.s.* system the surface tension is measured in dynes per centimetre. The dimensions of surface tension, are [Force]÷[Length] or $[MT^{-2}]$.

In the arrangement shown in Fig. 133, the two limbs AB, BC are not parallel, for if they were the arrangement would not be in stable equilibrium, but in neutral. For in this case (Fig. 134) the length l of the film in contact with the movable rod EF is constant, and hence the force $2Tl$ exerted by the film is independent of the position of the rod EF. Since the downward force Wg is also independent of the position of EF, if these two forces are exactly equal the rod EF will remain wherever it is put. If, however, we have not succeeded in exactly adjusting W to the right value, then if W is too small EF will be drawn up till it is in contact with BC, or if W is too great EF will fall till the ends A and D are reached, when the film will break. When the side wires are inclined as in Fig. 133, the length of the film in contact with DE, and hence the force exerted by surface tension varies with the position of the cross bar. If, when the bar is at DE, Wg is greater than $2Tl$, the bar will fall to some such position as D′E′; so that the new value of l, say l^1, exactly fulfils the condition $2Tl^1 = Wg$. If, on the other hand, W is too small, the bar will rise and l diminish till this relation is fulfilled.

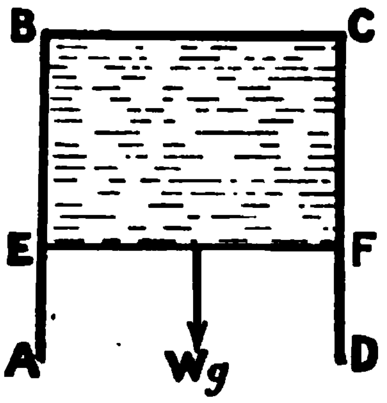


FIG. 134.

In the case of the arrangement shown in Fig. 134, if we start with EF in contact with BC, and then pull it down into the position shown, we shall in doing this have to do work, since we are moving EF against a force of $2T.\overline{\text{EF}}$. The work done is

$$2T.\overline{\text{EF}} \times \overline{\text{BE}},$$

since $\overline{\text{BE}}$ is the distance through which EF has been moved against the force. The energy corresponding to this work is stored up in the film, and may be recovered by allowing the film to contract. Hence if E is the energy of the film due to the surface tension, or the superficial energy, we have

$$E = 2T.\overline{\text{EF}} \times \overline{\text{BC}}.$$

But EF × $\overline{BC}$ is the area of the film A, say,

$$\therefore \qquad E = 2TA,$$

or

$$T = \tfrac{1}{2}\frac{E}{A}.$$

Hence, since E is the energy of the two surfaces of the film, each of area A, $E/2A$, or T, is the energy per unit of area of a single surface of the film. The dimensions of T obtained from this consideration are [Energy]÷[Area], or $[ML^2T^{-2}]\div[L^2]$, or $[MT^{-2}]$; the same result as that obtained before. The fact that a soap film possesses a store of potential energy is very evident when it breaks, for this potential energy immediately becomes kinetic energy, and the liquid of which the film was composed is projected with considerable velocity in all directions.

158*. Pressure within a Soap Bubble.—In a soap bubble, the pressure inside must be greater than the external pressure, on account of the surface tension of the film, which tends to make the bubble contract. Let the bubble be a sphere of radius R, and the pressure inside exceed the external pressure by a quantity p. Then, if we suppose the bubble divided into two hemispheres by a solid plane partition ABCD (Fig. 135), the area of this partition will be πR^2, and the downward force on the upper surface due to the excess of pressure p will be $\pi R^2 p$. The film meets the partition at right angles round the circumference of the circle ABCD, or along a length $2\pi R$. Hence if T is the surface tension of the liquid, the upward force exerted by the surface tension on the partition is $2T \times 2\pi R$, since the film has two surfaces, or $4\pi TR$. Hence, since this upward force must be equal to the downward force due to the excess pressure, we have—

FIG. 135.

$$\pi R^2 p = 4\pi TR,$$

or

$$p = \frac{4T}{R}.$$

From this expression we see that the pressure inside a soap-bubble decreases as the bubble gets larger. By measuring the pressure p within a bubble, and also measuring the radius R, the value of the surface tension T can be obtained.

159. Angle of Contact.—If a plate of glass is plunged in water with its side vertical, it will be found, as shown at (*a*), Fig. 136, that where the liquid touches the glass it is drawn up above the level of

the general surface. If, however, the glass is placed in mercury, the surface of the liquid near the glass is depressed below the general surface, as shown at (*b*). The angle BPA between the tangent to the liquid surface at the point P, where it meets the solid, and surface of the solid is called the ***angle of contact*** between the liquid and the solid. The angle of contact between a solid and a liquid depends on the third material, which exists above the free surface of the liquid. Thus the angle of contact between mercury and glass, when air is above the mercury, is different from the angle of contact when there is a layer of water above the mercury. In the case where the angle of contact APB is less than 90° ((*a*), Fig. 136), the surface tension of the liquid surface supports the part of the liquid which is above the general level. In the same way the surface tension, when the angle of contact is greater than 90° ((*b*), Fig. 136), withstands the hydrostatic pressure due to the liquid displaced near the surface of the solid.

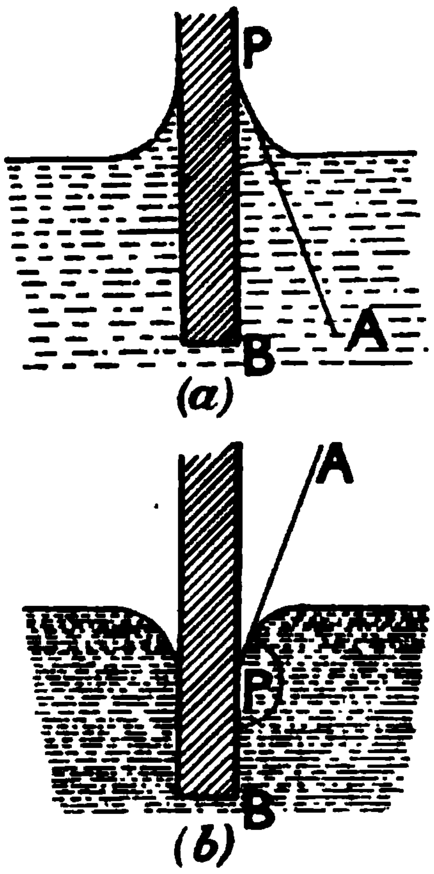


FIG. 136.

160. Capillarity.—If a clean glass tube of fine bore is dipped into water, the water rises inside the tube and stands at a level higher than the surface of the external water. This elevation of the water is due to the angle of contact between glass and water being less than 90°, so that the surface tension tends to raise up the water near the glass. Suppose a tube CB (Fig. 137) of radius r dips into a liquid, and the angle of contact between the solid and liquid is α. The surface of the liquid meets the solid along the circumference of a circle of radius r, and so, if the surface tension of the liquid is T, the force exerted by the tension of the surface on the inside of the tube is $2\pi rT$. This force is everywhere directed along the tangent to the liquid surface at the point of contact, *i.e.* along PA, and makes an angle α with the side of the tube. The walls of the tube will therefore react on the liquid with an equal and opposite force. The resolved part of this force parallel to the axis of the tube is $2\pi rT\cos\alpha$, and it is this vertical force which supports the column of liquid in the tube, so that the weight of liquid in the tube above the level of the general surface outside must be equal to this resolved force. If ρ is the density of the liquid, and h is the amount of the elevation in the tube,

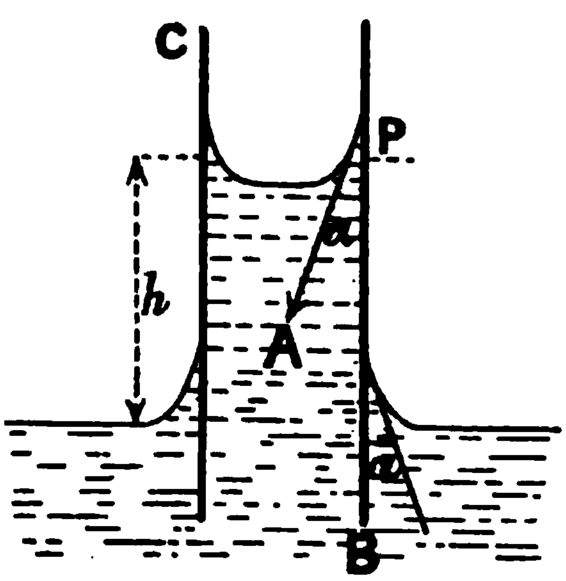


FIG. 137.

the volume of liquid raised is $\pi r^2 h$, and the mass raised is $\pi r^2 h\rho$. The downward force exerted by gravity on this mass is $\pi r^2 h\rho g$. Hence

$$2\pi r T \cos \alpha = \pi r^2 h\rho g,$$

or

$$h = \frac{2T \cos \alpha}{\rho g r}.$$

Thus for a tube of a given material and a given liquid, so that α, T, and ρ are constant, h is inversely proportional to r. If α is greater than 90°, $\cos \alpha$ is negative, and hence h is negative, that is, the level of the liquid inside the tube will be *below* the general surface.

If the liquid wets the walls of the tube, the angle of contact α is zero, so that $\cos \alpha = 1$, and

$$h = \frac{2T}{\rho g r},$$

or

$$T = \frac{h\rho r g}{2}.$$

Thus by measuring the radius of a capillary tube and the capillary elevation we can calculate the value of the surface tension T. Most of the accurate measurements of the surface tension of liquids have been made by means of capillary glass tubes, the capillary elevation being measured with a cathetometer, and the radius deduced from the weight of a thread of mercury filling a measured length of the tube.

161. Phenomena due to Surface Tension.—The apparent attractions or repulsions exhibited by small floating bodies on the surface of a liquid are due to surface tension. Thus two small pieces of wood floating on the surface of water rush together if they come within about a centimetre of one another. This is due to the fact that the angle of contact between water and wood is less than 90°, so that the water is slightly raised up between the two floating bodies as in a wide capillary tube. The pressure in the liquid between is less than in the surrounding mass, and the bodies come together. In the case where the angle of contact is greater than 90°, as, for instance, with needles floating on mercury or greased wood on water, the liquid between the bodies is depressed, and the hydrostatic pressure on the outside forces the bodies together.

If small fragments of camphor are placed on a clean-water surface, they dart about in a most life-like manner. This is due to the fact that the camphor dissolves slowly in the water, and that the surface tension of a solution of camphor in water is less than that of pure water. Hence, if the camphor dissolves a little faster at one side of the floating fragment than at the other side, the surface tension at the first side is reduced most, and the greater surface tension on the other side draws the fragment away.

162*. Viscosity.—If a liquid flows over a horizontal plate AB (Fig. 138), then the layer of liquid next the surface of the solid is, on account of cohesion, at rest, and the velocity of the fluid particles in the different layers is greater, the greater their distance from the solid. Hence the successive layers of liquid have different velocities, and as a result the more slowly moving layer tends to retard the motion of the adjacent more quickly moving layer, and is itself accelerated by the action of this layer. Thus any horizontal layer is acted upon above by a tangential force in the direction of motion of the liquid, and below by a second tangential force in an opposite direction. These two forces are due to what is called the viscosity of the liquid, which is really a kind of friction between the particles of a liquid when the different parts of the liquid are moving with different velocities. If a small surface C in the liquid of area a be taken, parallel to the fixed plane AB and at a distance s from it; then, if v is the velocity with which the liquid is moving at C, the tangential resistance (R) to the motion experienced by C is found to be given by the expression

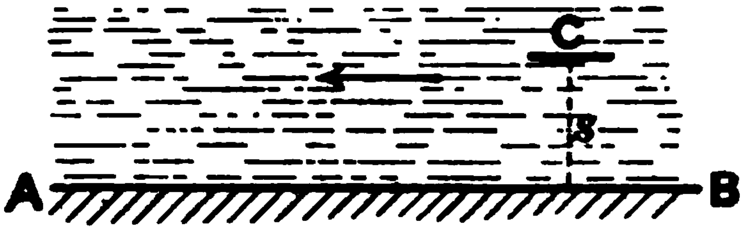


FIG. 138.

$$R=\frac{va}{s}\cdot\eta,$$

where η is a constant which depends on the nature of the liquid. If R, v, a, and s are all measured in *c.g.s.* units, then η is called the coefficient of viscosity of the liquid, and may be defined as the tangential force per unit area of either of two horizontal planes at the unit of distance apart, one of which is fixed, while the other moves with unit velocity, the space between being filled with the viscous liquid (Maxwell).

In the case of a liquid escaping by a long and narrow tube, the velocity of efflux depends not only on the difference of pressure (p) between the ends of the tube and the radius (r) of the tube, but, on account of viscosity, also on the length of the tube (l) and the coefficient of viscosity (η). The volume (V) of a liquid which escapes in one second is, according to Poiseuille, given by the equation

$$V=\frac{\pi p r^4}{8l\eta}.$$

Thus by measuring the volume of a liquid which escapes from a tube under a given difference of pressure, the coefficient of viscosity can be determined. The dimensions of the coefficient of viscosity are $[L^{-1}T^{-1}M]$. The following table gives the coefficient of viscosity in *c.g.s.* units for some liquids :—

COEFFICIENTS OF VISCOSITY.

Liquid.	Temperature. Deg. C.	Coefficient of Viscosity in *c.g.s.* Units.
Glycerine	2.8	42.20
"	20.0	8.30
Water	0.0	0.0178
"	10.0	0.0131
"	30.0	0.0081
Mercury	17.2	0.0160
Ethyl alcohol	10.0	0.0153

163. Solution.—In certain cases when a solid and liquid are mixed, the solid, or at any rate part of it, becomes dissolved in the liquid, forming a homogeneous liquid, the properties of which may differ considerably from those of the pure liquid. The solid and liquid parts of such a solution cannot be separated by mechanical means, such as filtration, nor will the solid separate out on allowing the solution to stand, although the density of the solid may be very different from that of the liquid (the solvent).

If we introduce a few crystals of common salt (sodium chloride) into some water at the ordinary temperature they will dissolve; but if we continue adding the salt to the same water, a time will come when no more of the salt will dissolve. Under these circumstances we are said to have a saturated solution. If the temperature of the saturated solution of common salt is raised, more salt will be dissolved; on allowing the temperature to fall to its original temperature, this additional salt will be deposited from the solution in the form of crystals. Hence at any temperature there is a fixed mass of a given solid which can be dissolved by unit mass of a given solvent to form a saturated solution. This mass of solid is called the coefficient of solubility of the solid in the given solvent at the given temperature.

In addition to having a solution of a solid in a liquid, we may have a solution of a liquid in another liquid. Here, however, we have to deal with two cases. We may have two liquids, such as alcohol and water, which mix, or dissolve, one in the other in all proportions, and are said to be miscible. Or we may have two liquids, such as ether and water, which are not soluble in all proportions. Thus a given mass of water will only dissolve a small quantity of ether, forming a saturated solution of ether in water; similarly, a given volume of ether will only dissolve a small quantity of water, again forming a saturated solution, but in this case of water in ether.

164. Diffusion of Liquids.—If two liquids which are miscible are introduced into a vessel so that the denser is below and the lighter is above, then, just as in the case of gases, diffusion will take place, some of the lighter liquid travelling down and mixing with the heavier liquid, and *vice versa*.

The rate at which liquids diffuse is, however, extremely small, as compared with the rate at which gases diffuse. This fact confirms the molecular theory of the constitution of liquids and gases, for in gases it is supposed that the molecules travel about, only occasionally coming near enough to other molecules to influence their motion, the greater part of their path being traversed uninfluenced by other molecules. In a liquid, on the other hand, although the molecules move about, they never get far enough away from the adjacent molecules to escape from the influence of these molecules, so that although a liquid molecule may move relatively to neighbouring molecules, as soon as it passes out of the range of influence of one set of molecules it comes within the range of other molecules. Hence the molecular motion in liquids is much more constrained than in the case of gases, and we should expect the rate of diffusion to be slower.

The rates at which liquids (in most cases solutions of salts) diffuse into water were experimentally determined by Graham in the following manner. A small wide-mouthed bottle A (Fig. 139), filled with the liquid, was closed by a glass plate, and then placed in a larger vessel B containing water, so that the surface of the water was above the top of the bottle A. The glass plate was then carefully slid off the top of the bottle, and the liquids left to diffuse. After a certain time samples of the different layers of the mixed liquids were drawn off by a pipette, and the composition of the solution determined. With solutions of the same substance of different strengths, Graham found that the rates of diffusion were proportional to the strengths of the solution. The rates of diffusion of different substances are, however, very different.

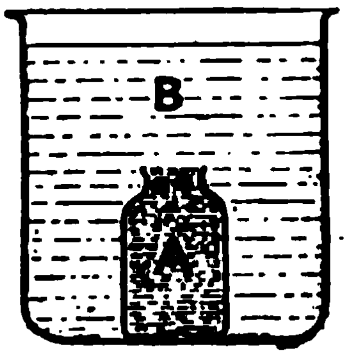


FIG. 139.

If we consider a small cylinder, one centimetre long and one square centimetre in cross section, and if the concentrations of the solution of a salt in water, *i.e.* the mass of the salt contained in unit volume of the solution, at the two ends of the cylinder differ by unity, then the quantity of salt in grams which will diffuse through the cross section of the cylinder, *i.e.* through unit area, in a *day* is called the diffusion constant of the salt. The following table gives some values of the diffusion constant in grams per square centimetre per day :—

DIFFUSION CONSTANTS.

Hydrochloric acid . . .	. 2.3	grams	per sq. cm.	per day.
Sodium chloride . . .	. 0.75	,,	,,	,,
Urea	. 0.81	,,	,,	,,
Cane sugar	. 0.31	,,	,,	,,
White of egg	. 0.05	,,	,,	,,
Caramel	. 0.02	,,	,,	,,

A consideration of the table of the diffusion constants given above shows that the rates of diffusion of different substances vary very considerably. Thus hydrochloric acid diffuses about one hundred times as fast as caramel. For this reason bodies have been subdivided into two classes, one containing such bodies as hydrochloric acid and the salts of the mineral acids, which are mostly crystalline, and diffuse comparatively rapidly. These are called crystalloids. The other, containing such bodies as gum, albumen, caramel, and the like, which are glue-like bodies of amorphous form that diffuse very slowly, and are called colloids.

A film of a colloid, such as paper coated with starch, if placed as a partition in a vessel, with pure water on one side and a solution of crystalloids and colloids on the other, will allow the crystalloids to diffuse through into the water, but entirely stops the passage of the colloids. Thus a colloid septum prevents the diffusion of other colloids, but allows the diffusion of crystalloids.

165. Osmosis.—If in a vessel A (Fig. 140), such as a thistle funnel with its larger end closed by a sheet of parchment, we place a solution of copper sulphate, filling the vessel up to about D, and then place it as shown in the figure, so that the parchment is below the surface of some pure water contained in a vessel C. Then it is found that the water makes its way through the parchment partition into A, the solution inside gradually rising up in the tube DB. Thus the water has been able to pass through the parchment in opposition to the hydrostatic pressure due to the column of liquid $\overline{BD}$. After a time the water ceases to force its way through the partition, its tendency to do so being counterbalanced by the hydrostatic pressure. It will also be noticed that in time some of the copper sulphate travels out into the surrounding water. If, instead of placing the vessel A containing the copper sulphate solution in pure water, it is placed in a solution of copper sulphate of the same strength as that inside, no change in the quantity of liquid in the vessel takes place. If, however, it is placed in a stronger solution, water will pass out from the vessel A, so that the solution inside becomes more concentrated. These phenomena are called osmosis, and the pressure produced in the vessel containing the salt solution, when placed in water, is called the osmotic pressure.

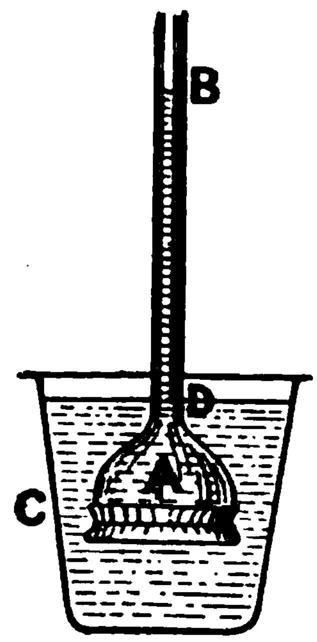


FIG. 140.

By using as the separating membrane a substance which, while it readily permits the passage of pure water, is impervious to the passage of certain substances when dissolved in the water, Pfeffer was able to measure the osmotic pressure due to solutions of different substances at different concentrations. Thus in the case of cane sugar such a

semi-permeable membrane is prepared by depositing ferrocyanide of copper within the pores of a porous earthenware cylinder. The cylinder, filled with the sugar solution, is plunged into pure water, and the maximum pressure developed inside measured by means of a manometer.

The osmotic pressure may be explained by supposing that the semi-permeable membrane is struck on both sides by the water molecules, but since there are fewer water molecules per unit volume inside, some of the space being occupied by sugar molecules which cannot traverse the membrane, more water molecules will in a given time strike the outside of the membrane than the inside, and hence, as the water molecules can pass through the membrane, more water molecules will enter than leave.

In the following table some of Pfeffer's results are given :—

Percentage of sugar in solution.	Osmotic Pressure in cm. of Mercury = P.	Density of Solution.	Mass of Sugar in 1 cc. of Solution.	Volume of Solution containing 1 grm. of Sugar = V.	PV.
1	53.5	1.0039	0.001004	99.6 c.c.	5329
2	101.6	1.0078	0.002016	49.6 "	5039
4	208.2	1.0158	0.004063	24.61 "	5124
6	307.5	1.0198	0.006119	16.34 "	5025

It will be seen that, if we consider the sugar when in solution as occupying the volume occupied by the solution, then the product of the osmotic pressure (P) into the volume (V) occupied by a gram of sugar is constant. This result, as was first pointed out by Van 't Hoff, corresponds to Boyle's law for gases.

Measurements of the change of osmotic pressure with temperature have shown another remarkable relation between the behaviour of a dilute solution and of a gas. Thus the osmotic pressure (P) of a 1 per cent. solution of cane sugar at a temperature t is given by the formula

$$P = 49.62(1 + 0.00367t).$$

The coefficient 0.00367 will be found later on to be the same as that for the variation of the pressure of a gas with temperature (§ 196).

It is only for *dilute* solutions that the above resemblances of the behaviour of the dissolved body and a gas hold. It would, however, appear that in such a dilute solution the molecules of the *dissolved* body exist in a condition in some way resembling that which occurs in a gas. We shall see later, particularly when we come to consider the electrical properties of dilute solutions, what suppositions have been made to account for the fact that it is only when dilute that the solutions obey the above gaseous laws.

CHAPTER XIX

PROPERTIES OF SOLIDS

166. Isotropic Bodies.—A body in which a spherical portion, when tested in different directions, exhibits no difference in its physical properties is said to be *isotropic*. Except under very special conditions, all liquids and gases are isotropic. Some solids, however (for instance, crystals), exhibit different physical properties in different directions, and are called ælotropic. In most of the following sections we shall deal exclusively with the properties of isotropic solid bodies.

167. A Perfect Solid.—When discussing the distinction between solids and liquids, we pointed out that there was no clear line of demarcation, but that from a rigid solid, such as glass, there is a continuous series extending through soft solids such as lead and butter, very viscous liquids such as sealing-wax and pitch, to treacle and glycerine. Just as in considering the behaviour of liquids we dealt with a typical liquid such as water, so in the case of solids we shall consider one in which, after suffering a strain which alters its *shape*, on the removal of the stress it completely regains its former shape. Such a solid is called a perfect solid, and the above conditions are practically satisfied by many solid bodies so long as the deforming stress does not surpass a certain value.

168. Malleability and Ductility.—By malleability is meant the property possessed by some solids of being beaten into thin sheets without losing their continuity. Of all materials pure gold possesses the property of malleability to the most marked degree. Thus, when preparing gold-leaf, a piece of gold is first rolled into a sheet somewhat thinner than foreign note-paper, next a portion is beaten out between two sheets of vellum till its surface has been increased, and therefore its thickness decreased about twenty-fold. This twenty-fold decrease of thickness, without rupturing the sheet, can be again twice repeated.

By ductility is meant the property of being drawn out into fine wires. A rod of the metal is passed in succession through a number of holes, each a little smaller than the last, the diameter of the rod continually decreasing, while its length is correspondingly increased.

169. Hardness.—When one body can be made to scratch a second, but cannot be scratched by it, we say that the former body is harder than the latter. Although some attempts have been made to devise a means of accurately measuring the hardness of bodies, they have not been

attended with much success. All that can be done at present is to give a body's position, as far as hardness is concerned, in a scale of hardness composed of various bodies. The scale usually adopted, and due to Mohs, is as follows, the first being the softest :—1. Talc ; 2. Crystallised Gypsum ; 3. Calcspar ; 4. Fluorspar ; 5. Apatite ; 6. Felspar ; 7. Quartz ; 8. Topaz ; 9. Sapphire ; 10. Diamond. A body having a hardness of 6.5 would be one which would scratch felspar, and be scratched with about the same ease by quartz.

170. Elasticity of Volume.—Solids, with few exceptions, are very slightly compressible, in this property resembling liquids. The volume elasticity of a solid is measured in the same way as that of a liquid. Thus if a uniform pressure of p dynes per square centimetre, acting everywhere normal to the surface of the solid, such as would be produced if the solid were immersed in a liquid under a pressure p, is applied, and the volume changes from V to $V-v$, then the coefficient of compressibility, or the volume elasticity, of the solid is $p \div v/V$ or pV/v. The following table gives the value of the volume elasticity of some solids :—

VOLUME ELASTICITY OF SOLIDS.

Glass	4.1×10^{11}	dynes/cm^2.
Brass	10.6	„
Iron (wrought)	14.9	„
Steel	18.8	„

171. Elasticity of Shape (Rigidity).—The elasticity of shape or rigidity of a solid is measured by the ratio of the stress, *i.e.* the force producing the change of shape, to the strain, *i.e.* the change in shape, produced. The shape of a body may be altered in various ways : thus if weights are attached to one end of a wire, the other end being held fast, the wire stretches ; on the removal of the weights, so long as the wire has not been too much deformed, it regains its original length. Another way of altering the shape of a body is to twist one end while the other end is held fast ; or again, if one end of a rod is held in a vice, and the other end pulled on one side, the rod becomes bent ; in each case the elasticity of the solid will resist the deformation, and when the stress is removed will cause the body to resume its unstrained position. We shall consider each of the above methods of straining a solid separately.

172. Elongation : Young's Modulus—Hooke's Law.—Suppose a wire of length L and radius r, when stretched by a force P in the direction of its length, increases in length by an amount l, then the stress, or force per unit area, acting on the wire and tending to increase its length is $P \div$ area of cross section, or $P/\pi r^2$. The total elongation being l, the strain, or elongation per unit length of the wire, is l/L. Hence the modulus of elasticity, Y, is given by

$$Y = \frac{\text{stress}}{\text{strain}} = P/\pi r^2 \div l/L = \frac{PL}{\pi r^2 l}.$$

Y is called Young's modulus, and it may be experimentally determined by means of the arrangement shown in Fig. 141. Two wires of the material to be tested are securely fastened to an overhead beam at A. To one of these wires is attached a small, finely divided scale B, and to the other a vernier. Attached to the lower end of one wire are two weights D, which serve to keep the wire stretched tight, and to the lower end of the other wire is attached a scale-pan E, in which the weights used to stretch the wire can be placed. The elongation produced by the weights is measured by the vernier and scale. The object of the second wire is to eliminate the effects of any change in length produced in the wire by a change of temperature, since such a change would affect both wires to the same extent, and hence would not affect the reading on the scale. The same remark applies to any give of the support at A produced by the added weights. So long as the weight used to stretch the wire is not so great as to produce a permanent elongation of the wire, it is found that the elongation is proportional to the stretching force. This is known as Hooke's law.

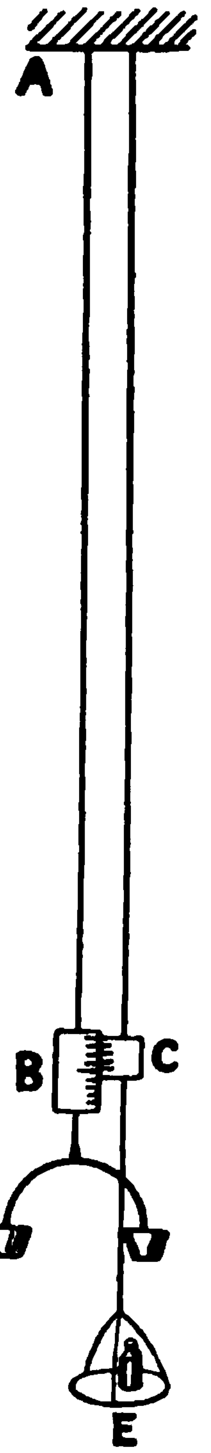


FIG. 141.

The following table gives the value of Young's modulus for some metals:—

YOUNG'S MODULUS.

Steel (hard drawn) . . .	$2\cdot0 \times 10^{12}$	dynes/cm^2.
Iron (wrought)	1.9	„
Silver (drawn) . . .	0.7	„
Brass (drawn)	1.1	„

173. Bending.—When a rod AB (Fig. 142), firmly fixed at A, has a force applied at B at right angles to AB, it becomes bent into such a form as AB′. In this case the upper parts of the rod have been stretched, while the lower parts have been compressed, so that, except for a thin band down the middle, the strain is really one of elongation. If the rod is rectangular in section, and of depth d and breadth b, the length being l, and a force of P dynes deflects the end through a distance l, then Young's modulus Y is given by

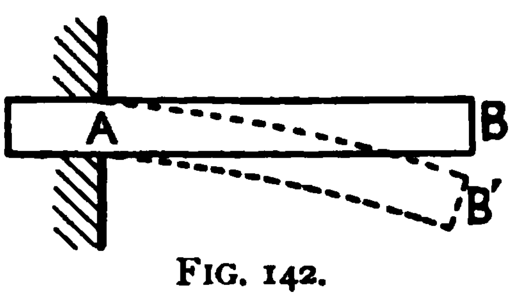


FIG. 142.

$$Y = \frac{4PL^3}{bd^3l}.$$

If, instead of being fixed at one end, the two ends of the rod are free, but are supported on two knife-edges placed at a distance L_1 apart, and l_1 is the distance through which the centre of the rod is deflected when loaded with a force P, Young's modulus is given by

$$Y=\frac{P}{4bd^3}\cdot\frac{L_1^3}{l_1}.$$

It will be noticed that in this case we are practically dealing with two rods each of length $L_1/2$, fixed as in the first case, and each acted upon by a force $P/2$ in the upward direction at the point where the rod rests on the knife-edges.

174. Torsional Rigidity.—If one end of a cylindrical wire of radius r and length l is kept fixed while a twisting couple u is applied to the other end, and under this twisting stress the end of the wire turns through an angle ϕ, it is found that so long as ϕ is not too great, it is proportional to the applied couple u, so that if the couple is doubled, the angle through which the end of the wire is twisted is also doubled. The value of ϕ, in terms of the dimensions of the wire, is given by the equation

$$\phi=\frac{2lu}{\pi r^4}\cdot\frac{1}{n},$$

where n is a constant depending on the nature of the material of the wire, and is called the simple rigidity or coefficient of torsional rigidity of the wire. It will be noticed, since ϕ is inversely proportional to the fourth power of the radius of the wire, that the deflection produced by a given couple increases very rapidly as the radius of the wire decreases. Thus if the radius of the wire is reduced to a half, the value of ϕ, corresponding to the same value of the deflecting couple, increases sixteen-fold. The importance of this rapid decrease of the torsional rigidity of a wire, when the diameter is reduced, comes in when we use the rigidity of such a wire to measure small forces and couples, as in the Cavendish experiment. By very rapidly drawing out a small stick of quartz, raised to a white heat in an oxy-hydrogen blowpipe, Boys has produced threads of fused quartz of such extreme fineness that a force of one dyne acting at the end of a lever 1 centimetre long (*i.e.* a unit couple) will twist one end of a fibre 10 centimetres long through 360°.

The following table gives the coefficient of torsional rigidity for some solids :—

SIMPLE RIGIDITY.

Steel	8.2×10^{11} dynes/cm^2
Iron (wrought)	7.7 ,,
Brass	3.8 ,,
Quartz (fused)	2.9 ,,
Glass	2.4 ,,

175*. Torsion Pendulum.—If a solid body, suspended by a wire, be twisted away from its position of rest and then released, it will execute S.H. vibrations about its position of rest, for the torsional rigidity of the wire will give a force tending to restore the body to its original position proportional to the deflection. If u is the restoring couple due to the rigidity of the wire produced when the body is twisted through unit angle (a radian), and K is the moment of inertia of the solid, then the time of oscillation is given by

$$t=2\pi\sqrt{\frac{K}{u}},$$

or, substituting the value of u in terms of the simple rigidity and dimensions of the wire,

$$t=r^2\sqrt{\frac{8\pi lK}{n}}.$$

Such a torsional pendulum can be used to prove that Hooke's law holds for torsional strains, that is, that the restoring couple or stress is proportional to the strain or twist, for the time of oscillation is found to be independent of the amplitude of the vibrations, and it is only when the restoring force is proportional to the deflection that this isochrony is secured.

176. Elastic Limit, Elastic Fatigue.—It is found that if a solid is deformed more than a certain amount, then, on the removal of the deforming stress, it does not completely regain its original form. Under these circumstances the body is said to have been strained beyond its elastic limit. The limits within which they may be considered as completely elastic vary very much with different materials. Thus quartz, and to a less extent steel and glass, can suffer a considerable strain, and yet when the stress is removed they will recover their original form; while soft iron, copper, and lead exhibit a permanent deformation or "set" even with quite small strains.

It is found that if the deforming stress is continued for a long time the strain produced gradually increases. This phenomenon is referred to as elastic fatigue, and it seems to show that for long-continued stresses the molecules even of solids gradually take up new configurations. A somewhat similar phenomenon is the fact that after a solid has been strained even below its elastic limit it does not, on the removal of the deforming force, immediately return completely to its original form, but only does so after some time. Thus if a silver wire is twisted in one direction and kept twisted for a day, and then twisted in the opposite direction for an hour, on being released it does not completely recover, but remains slightly twisted in the direction of the last twist. This residual twist gradually disappears, and then a slight twist in the direction of the first one appears, reaches a maximum, and then dies out. The wire thus "remembers" the deformation previously applied, and the residual effects appear in the opposite order to the original deformations.

BOOK II

HEAT

CHAPTER I

THERMOMETRY AND EXPANSION BY HEAT

177. Temperature.—Although we are able in many cases to distinguish by our sensations between hot and cold bodies—for instance, we can by touch often determine which of two bodies is the hotter—yet our senses do not permit of our forming a *quantitative* estimate of the amount by which one is hotter than the other. In ordinary language we use the words hot, warm, tepid, cool, cold, &c., to indicate a series of states of a body with reference to heat. In scientific language we use the word *temperature* to express the same series of condition. Thus a hot body is said to have a higher temperature than a cold body.

As we shall see in the following pages, the characteristic which above all others distinguishes bodies of which the temperatures differ is, that if these bodies are placed in contact, then heat will of itself pass from the one to the other until they reach the same temperature. That body which loses heat during the process of equalisation is said to be at the higher temperature.

It is found that not only does the sensation we experience when we touch a body vary with the temperature, but also that most of the physical properties of matter change when the temperature changes. Thus the density, elasticity, refractive index, &c., of a body all depend on the temperature.

In order to have a means of measuring temperature, we make use of the change in some physical property of some kind of matter which takes place as the temperature of the body changes. The physical property which is most often employed for this purpose is the length of a solid, or the volume of a liquid or gas, both of which depend on temperature. In order to define certain fixed temperatures, we also make use of the fact that the physical state of a body depends on the temperature. Thus according to the temperature we may have the same kind of matter existing as a solid, a liquid, or a gas, as, for instance, ice, water, and

steam. It is found that during the time the change from one state to the other is going on the temperature remains constant. Thus if a quantity of pounded ice is heated over a flame, the whole being kept well stirred, although the ice becomes gradually converted into water the temperature does not rise till the last particle of ice has been melted, the heat supplied by the flame being simply used up in changing the body from the solid state into the liquid state. If, after the ice is all melted, the heating is continued, the water will eventually begin to boil, becoming converted into the gaseous state (steam), and during the change the temperature of the remaining water will remain constant. It will thus be seen that we may use the temperature at which a given substance, under given conditions, changes its state as fixed points on a scale of temperature.

In order to subdivide the interval between these two temperatures, use is made of the change in volume of some fluid, usually mercury or hydrogen, which occurs with change in temperature. Now there is no *à priori* reason for supposing that the *rate* of change of volume of a substance, say mercury, with temperature is the same at all temperatures. Since, however, we have no special means of measuring temperature as distinct from the effects of temperature on the physical properties of bodies, we have, at any rate as a starting-point, to *assume* that the rate of change of some fixed property of some standard substance is constant, and to use this change to subdivide the temperature between our two fixed points. For the present, at any rate, we shall take the change in volume of mercury when contained in a glass vessel as the means of defining the temperature between our fixed points.

178. Thermometric Scales.—The lower fixed point of most scales of temperature is the temperature of melting ice under ordinary atmospheric pressure. The upper fixed point is the temperature of the steam given off from water boiling under the pressure of one standard atmosphere (§ 133).

Let v be the apparent increase in volume of a given mass of mercury enclosed in a glass envelope when its temperature is raised from that of melting ice to that of water boiling under standard conditions. Then the interval of temperature which will cause this quantity of mercury to expand by an amount $v/100$ is called a degree Centigrade, and is indicated by the symbol, 1° C. On the Centigrade scale (first used by Celsius) the temperature of melting ice is called zero (0° C.), and that of boiling water 100° C., the interval, as has been said, being divided into a hundred degrees. For temperatures below that of melting ice the scale is continued downwards, the sign minus being prefixed. Thus a temperature such that the volume of the above mass of mercury is $\frac{5v}{100}$ *less* than at 0° C. is indicated by $-5°$ C. In a similar way the scale is continued above 100° C. In all scientific work, and, with one or two

exceptions, for everyday use also, the Centigrade scale of temperature is employed in all countries. There are, however, two other scales of temperature occasionally employed. In one of these, the Fahrenheit scale, the temperature of melting ice is called 32° F., and that of water boiling under standard conditions 212° F., the interval being divided into 180°.

The Fahrenheit scale is in use in England for commercial and meteorological purposes. The other temperature scale is one due to Réaumur, on which the temperature of melting ice is called 0° R., and that of boiling water 80° R., the interval being divided into 80°.

Any reading on one of these scales can easily be converted to the corresponding reading on either of the others, for from the definitions 100 degrees on the Centigrade scale are equal to 180 degrees on the Fahrenheit scale, and to 80 degrees on the Réaumur scale. Let c, f, and r indicate the readings corresponding to the *same* temperature on the three scales. Then since 0° C. = 32° F. = 0° R., if we deduct 32° from the Fahrenheit reading we have the number of degrees on each scale by which the given temperature differs from the temperature of melting ice.

Hence

$$c : f-32 : r :: 100 : 180 : 80,$$

or

$$\frac{c}{100} = \frac{f-32}{180} = \frac{r}{80}.$$

By means of these equations, any temperature on any one scale can be immediately converted into either of the other scales.

179. The Mercury Thermometer.—The most commonly employed instrument for the measurement of temperature is the mercury thermometer. The ordinary form of a mercury thermometer is shown in Fig. 143. It consists of a glass bulb B containing mercury, connected

FIG. 143.
(*From Watson's "Elementary Practical Physics."*)

to a stem CA, which is traversed by a fine, uniform capillary bore. Such a thermometer is filled by attaching a small funnel to the open end of the tube, and placing some pure and dry mercury in it. The bulb is then slightly heated, some of the contained air being thus expelled, and, on cooling, a little mercury is driven into the bulb by the atmospheric pressure. This mercury is then heated till it boils, the mercury vapour driving out the air, so that when the bulb cools the mercury is driven in and completely fills the bulb and tube. The thermometer is then heated up to the highest temperature which it is intended to measure, and the end of the tube closed in the blowpipe flame. On cooling, the mercury contracts and the column sinks in the stem, leaving a vacuum

at the top of the tube. In order to prevent the mercury, at high temperatures, distilling to the top of the bore, the tube above the mercury is sometimes filled with some gas which does not act chemically on mercury, such as nitrogen, under pressure.

In thermometers intended for accurate measurements, the scale, as is shown in Fig. 143, is generally engraved on the outside of the stem. In some German thermometers the stem is made very narrow, and the scale is engraved on a separate piece of opal glass attached to the back of the stem, the whole being enclosed in a glass tube, in order to protect the scale and tube, and so that their relative position should not vary.

180. Determination of the Fixed Points of a Thermometer.—The freezing-point of a mercury thermometer is determined by surrounding the bulb and the stem up to the zero mark with pure snow, or finely pounded pure ice, then pouring over the snow or ice some distilled water. The thermometer is allowed to stand in the mixture of ice and water till the reading becomes constant. It is of utmost importance that the ice or snow used in determining the freezing-point should be quite free from contamination, such as salt, as otherwise the zero obtained will be too low (§ 225).

In determining the upper fixed point (the boiling-point) the thermometer is suspended in the arrangement shown in Fig. 144, so that the end of the mercury column just projects above the top. The steam rising from boiling water in the vessel A passes up the tube B into the inside tube C, which surrounds the thermometer, then down between the inside and outside tubes, and escapes by a lateral opening E into a Liebig condenser F. Here, by means of a stream of cold water which is passed round the outside, it becomes condensed and returns as water to the boiler through the tube G. The barometric height must be noted, for the temperature of the stream rising from boiling water varies with the pressure. It is only when the pressure corresponds to the weight of a column of mercury 76 cm. high, measured

FIG. 144.

at sea-level and at latitude 45°, that the temperature is 100° C. If the pressure is different a correction will have to be applied, which may be obtained from tables giving the temperature of the steam given off from boiling water at different pressures (§ 218).

The reason for placing the thermometer in the steam and not in the boiling water is because it is found that the temperature of the boiling *water* depends to a certain extent on the nature of the vessel in which the water is contained. Another important consideration is that small quantities of impurity in the water alter the temperature of the water at which ebullition takes place, but do not affect the temperature of the *steam*.

181. Calibration of the Thermometer Tube.—If the bore of a thermometer were of exactly uniform cross section throughout its whole length, equal lengths of the bore would everywhere have equal capacities, so that if the stem between the two fixed points were divided into 100 equal parts, the volume of the bore between any two divisions would be exactly $\frac{1}{100}$ of the volume between the fixed points. In practice, however, this condition is never quite accurately fulfilled, the bore of the tube varying slightly in cross section from point to point, so that equal lengths no longer represent equal volumes. Since we use equal increments of volume of the mercury to measure equal increments of temperature, it becomes necessary either to place the divisions of the scale, not at equal distances apart, but so spaced that the volume of the bore between any two divisions is everywhere the same, *i.e.* make the divisions closer together in those parts where the bore is wide, and further apart where the bore is more narrow, or, having divided the stem into divisions of equal length, to determine a series of corrections to be applied to the readings to allow for the inequalities of the bore.

To subdivide a tube into divisions corresponding to equal volumes, or calibrate the tube, as it is called, a thread of mercury 2 or 3 cm. long is drawn into the tube, if the calibration is performed before the thermometer is made, or, in the case of a finished thermometer, a short length of the mercury column is separated from the remainder either by jerking, or by heating the column at the point where separation is required by means of a small gas jet about 3 mm. high. This thread is moved to different parts of the tube, and its length measured either by means of the graduations on the tube or by a horizontal cathetometer. Then, since each of these lengths corresponds to a volume of the bore equal to that of the thread of mercury, we get the tube divided into intervals of equal volume. By using threads of various lengths, a tube can in this way be very accurately calibrated. The process is, however, a very lengthy one, and the corrections of a thermometer are generally obtained by comparing its readings with those of a standard thermometer, which has been previously calibrated, when the two thermometers are placed in an enclosure so that they are at the same temperature.

182. Errors of Mercury Thermometers.—Two mercury thermometers, of which the tubes have been accurately calibrated, will give readings which agree very closely with one another if the glass of which they are composed is of exactly the same kind. Thermometers made of different kinds of glass do not, however, agree completely, owing to the fact that the different kinds of glass do not expand exactly alike, and that what we observe is the difference between the expansion of the mercury and of the glass envelope.

Another effect due to the glass is a gradual rise of the zero point,[1] which goes on a long time after the thermometer is made. This rise is rapid for the first few months; it then gradually becomes slower, but does not stop even after many years. The rise is due to the gradual recovery of the glass from the effects of the extreme heating to which it was subjected when the thermometer was made. The magnitude of this secular rise of the zero, as it is called, depends on the nature of the glass, and the following table gives some values for two kinds of glass, from which the magnitude of the effect can be judged :—

SECULAR CHANGE OF ZERO.

"Verre Dur."		"Crystal."	
Date.	Zero Reading.	Date.	Zero Reading.
Feb. 1885 .	−0.144	May 1885 .	+0.130
July. 1885 .	−0.121	Nov. 1885 .	+0.281
Nov. 1885 .	−0.112	March 1886 .	+0.323
March 1886 .	−0.106	May 1886 .	+0.376
Dec. 1886 .	−0.098	Oct. 1887 .	+0.501
Sep. 1887 .	−0.092		
Aug. 1888 .	−0.090		

It has been proposed to hasten the secular rise by maintaining the thermometer at as high a temperature as it will stand for some days, then to keep it for some days at a somewhat lower temperature, and so on, till the temperatures at which the thermometer is to be used are reached; the reason for this treatment being, that it is found that the secular rise is more rapid at high temperatures than at low, so that in this way the greater part of the rise can be got over in a few weeks.

In addition to the secular rise of the zero, a temporary lowering of the zero takes place when the thermometer has been heated even to comparatively *low temperatures*, due to the temporary enlargement of the bulb.

[1] Of course the rise in the zero reading is accompanied by an equal rise in all the other readings of the thermometer. The term "rise in the zero" is employed because this rise is generally detected by observations on the lower fixed point.

183. Maximum and Minimum Thermometers.—Thermometers for registering the highest or lowest temperature reached during any interval are much used in meteorology. The commonest forms are Rutherford's and Six's.

Rutherford's *maximum* thermometer consists of an ordinary mercury thermometer with a small iron index introduced into the bore. Since mercury does not wet iron, when the temperature rises the mercury column pushes the index before it, but when the temperature falls the mercury does not draw the index back, so that the end of the index next the bulb indicates the highest temperature reached. The *minimum* thermometer has alcohol for the thermometric liquid, and the index is of glass. The glass is wetted by the alcohol, and hence, when the temperature falls, the index is carried along by the retreating surface of the alcohol on account of capillarity. When the temperature rises, the alcohol flows past the index, but does not move it, and hence the end furthest away from the bulb indicates the lowest temperature reached.

FIG. 145.

In Six's thermometer the maximum and minimum temperatures are shown by two small indexes on the same thermometer. The thermometric liquid is alcohol contained in the bulb A, Fig. 145. At the end of the alcohol column is placed a thread of mercury BC, while the remainder of the tube and part of the bulb D is filled with alcohol. The two ends of the mercury thread serve to indicate the temperature. Two small glass rods, fitted with a small steel spring, as shown at G, are pushed up by the mercury but are kept from slipping down, when the mercury retreats, by the springs. The bottom of the index E shows the lowest temperature reached, since it indicates the smallest volume which has been occupied by the alcohol in A. In the same way the bottom of F shows the highest temperatures reached. The indexes can be moved down, and hence reset, by a small magnet which attracts the steel springs.

184. Linear Expansion of Solids.—If the temperature of a solid body is raised, the distance between any two points in the body in general increases, and the body is said to expand. Thus a cylindrical bar of a metal, say iron, when its temperature is raised, increases in length; it also increases in radius. If we only consider the increase in length that takes place in any given direction, we are said to deal with the linear expansion of the body.

This expansion of solid bodies with heat is made use of in many ways. Thus the iron tyres of carriage wheels are made slightly smaller

than the wooden rim. The tyre is then heated, so that it expands, and is placed round the wheel. On cooling the iron contracts, and binds the wood firmly together. The same kind of operation is gone through when making large guns, which are built up of a number of cylinders fitting one over the other. The inside of each cylinder is made just a little smaller than the outside of the previous one, then, by heating the outer, it expands till it will slip over the inner, and on cooling shrinks on to, and becomes firmly attached to the inner cylinder.

In all large metal structures, such as bridges, very careful provision has to be made so as to allow for the expansion and contraction which takes place with change of temperature, otherwise the structure would be strained or even ruptured.

If, when a bar of a solid of length l is heated from 0° C. to 1°, it increases in length to $l+\delta l$, then $\delta l/l$, or the increment of length of unit length of the bar, is called the coefficient of linear expansion. If a bar of which the coefficient of linear expansion is a is heated from, say, t_1 to t_2, and if L_0 is the length at 0°, then the length at t_1 will be $L_0+L_0at_1$, or $L_0(1+at_1)$, and the length at t_2 will be $L_0(1+at_2)$. Hence the increase in length, when heated from t_1 to t_2, is

$$L_0(1+at_2)-L_0(1+at_1),$$

or

$$L_0a(t_2-t_1)$$

If L_{t_1} is the length at t_1, since $L_{t_1}=L_0(1+at_1)$, the increase in length at t_2 is

$$aL_{t_1}(t_2-t_1)/(1+at_1),$$

or

$$aL_{t_1}(t_2-t_1)(1-at_1+\&c.),$$

since $\frac{1}{1+at_1}=1-at_1+$ terms in a^2 and higher powers of a.

Hence the increase in length is $aL_{t_1}(t_2-t_1)-L_{t_1}t_1(t_2-t_1)a^2$. If, therefore, as is the case, a is very small, so that terms in a^2 can be neglected on account of their extreme smallness, this reduces to

$$aL_{t_1}(t_2-t_1).$$

We therefore get, in the case of linear expansion, that the increase in length is equal to the original length multiplied by the coefficient of linear expansion and by the increase in temperature.

In the case of the expansion (cubical) of gases, the coefficient of expansion is so great that the term involving its square cannot be neglected, and hence the coefficient of expansion has to be defined, with reference to the state, at 0°. In the case of the linear expansion of solids, although for the sake of uniformity we have taken the increment in the length between 0° and 1° in the definition, the above investigation

shows that as a is always very small we need not have done so, but may simply define the coefficient of linear expansion of a solid as the increase in length of unit length, when the temperature is raised 1° C.

185. Measurement of the Coefficient of Linear Expansion.—Since the distance between the marks on all standards of length is only what it purports to be at one temperature, it is very important to know the coefficient of linear expansion of the material of which the standards are made, in order that we may be able to reduce measurements made at any temperature to what they would be if the standard had been at the temperature at which it is correct.

The amount by which solids expand when heated through any reasonable range of temperature being very small, a bar of iron one metre long at 0° C. expanding to 1.00117 metres when heated to 100° C., the chief difficulty in determining the coefficient of expansion lies in measuring the increase in length. Two methods have been used for this purpose. In the first of these, by a system of levers (partly material and partly optical) the expansion is greatly magnified, and then, from the relative lengths of the arms of the levers, the actual expansion is calculated from the magnified movement. In the other method the utmost refinement is introduced into the instruments for measuring small lengths, and the expansion is measured directly. The relative accuracy of these two methods is very different, since in the first method it is almost impossible to measure with any great degree of accuracy the actual amplification given by the levers. In the second method, however, by the combination of a microscope with a micrometer screw, or by means of optical interference, it is possible to measure the expansion with considerable accuracy.

As an example of the first method, we may take the experiments of Lavoisier and Laplace.

The bar HK (Fig. 146) to be measured is placed in a water bath, the end K resting against a firm stop. A lever D pivoted at G carries a

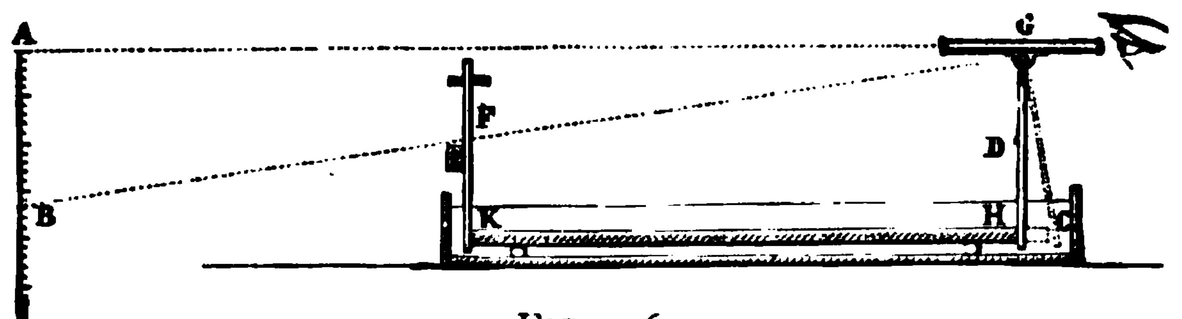


FIG. 146.
(From Ganot's "Physics.")

telescope at its upper end and rests against the end H of the bar. The telescope is focussed on a vertical scale AB fixed at some distance away. The bar being packed round with water and ice, the reading on the scale which coincides with the cross wire of the telescope is taken. The water

in the bath is then heated to a temperature t, which is measured by a thermometer placed alongside the bar, and the bar expands to C, the end K remaining against the stop. The expansion of the bar moves the lever, which tilts the telescope down, so that the image of the point B is now seen to coincide with the cross wire. Since the two triangles HCG, ABG are similar, we have

$$\frac{HC}{HG}=\frac{\overline{AB}}{\overline{AG}},$$

or

$$\overline{HC}=\overline{AB}\cdot\frac{\overline{HG}}{AG}.$$

If $\overline{HG}$ and $\overline{AG}$ are measured, the expansion $\overline{HC}$, and hence the coefficient of expansion of the bar, can be calculated from the difference, $\overline{AB}$, in the scale readings.

As an example of the second method, we may take the measurements made by Roy and Ramsden when determining the coefficient of expansion of the rods used in measuring an arc of the meridian.

Roy and Ramsden employed three troughs, AB, CD, and EF (Fig. 147). These troughs were of such a length that they could hold the bar to be experimented upon and two similar bars of iron. At either end of each of these bars an upright was fixed. The two uprights on the bar in AB carry the eyepieces of two small telescopes, these eyepieces being fitted with cross wires. The uprights on the bar in CD carry the object-glasses of the telescopes, while those on the bar in EF carry two fine spider lines forming a cross. The upright G on the bar in CD can be moved through a small distance parallel to the length of the bar by means of an adjusting screw. The central bar rests on rollers and is held by a spring which presses against the end nearer C, so that the other end rests against the point of a micrometer screw M. The three troughs are first filled with pounded ice or snow, so that the temperature of all three bars is 0° C.

The bar in CD is then moved parallel to its length by the micrometer screw till the image of the cross wires K, formed by the object glass L, coincides with the cross wires in the eyepiece N. The upright G is then moved till the image of the cross wires H coincides with the cross wires in the eyepiece O.

The ice in the trough CD is then melted and the water heated and well stirred, the temperature being read by means of thermometers placed alongside the bar. The bar expands and moves towards C, since the point of the micrometer screw prevents it moving in the opposite direction. The troughs AB and EF are kept full of ice, so that the bars they contain remain at 0°, and thus the distances between the uprights they carry remain unaltered. If, on account of the expansion of the bar between the upright L and the end on which the micrometer screw presses, the image of the cross wire K does not exactly coincide with the cross wire in the eyepiece N, the micrometer screw is turned, so that the bar is moved as a whole, till this coincidence takes place. The position of the micrometer screw is then read by means of a scale giving the whole turns, and a divided head by which the fractions of a turn are obtained. The micrometer screw is now turned, so that the bar is moved from C towards D, till the cross wires on H coincide with the cross wires in the eyepiece O, and its position is again read. The difference in the micrometer readings gives the amount by which the bar has expanded, and hence by measuring the distance between the uprights, *i.e.* the length of the part of the bar whose expansion has been measured, and knowing the interval of temperature through which the bar has been heated, the coefficient of expansion can be calculated.

More recent measurements of the coefficient of linear expansion have been made with the comparator described on p. 19; the distance between two fine marks on a bar of the metal, almost exactly one metre apart, being measured at two known temperatures.

The method of measuring the coefficient of expansion of small lengths of a solid by using optical interference bands will be described when dealing with light.

The following table gives the coefficient of linear expansion of some materials per degree Centigrade :—

COEFFICIENTS OF LINEAR EXPANSION.

Platinum	$.0899 \times 10^{-4}$
Copper	$.1678 \times 10^{-4}$
Steel (annealed)	$.1095 \times 10^{-4}$
Zinc	$.2918 \times 10^{-4}$
Brass	$.187 \times 10^{-4}$
Glass	$.083 \times 10^{-4}$
Pine } Parallel to fibre	$.054 \times 10^{-4}$
Ash } Parallel to fibre	$.095 \times 10^{-4}$

186. The Compensation of Timekeepers for Variation in Temperature.—A problem of considerable practical importance is to design a pendulum the length of which shall be the same at all temperatures, for, as we have seen in § 113, the time of vibration of a pendulum depends on its length. The result is that in an ordinary pendulum the length increases with increasing temperature, thus the clock goes slower. The problem to be solved is to design a pendulum-rod in such a way that the distance between the point of support and the centre of oscillation (practically the centre of the bob) shall be the same at all temperatures. One solution of this problem is to employ a rod composed of two metals having different coefficients of expansion. Let the coefficient of linear expansion of one of the metals, A, be α, and that of the other metal, B, be β. Suppose we have three rods of the metal A and two of the metal B, and that they are connected together as shown in Fig. 148, where the dark lines AB, CD, and EF represent the metal A, and the light lines BC, DE represent the metal B. Let the total length of the rods of the metal A be L', and the total length of the rods of the metal B be L'', both at a temperature of 0°. If now the rods made of A were *alone* heated to a temperature t, the end F would *descend*, supposing the end A fixed, through a distance

$$L'\alpha t.$$

If, on the other hand, the rods made of B had been alone heated to t, the end F would have *risen* through a distance

$$L''\beta t.$$

A

Fig. 148.

The condition that the descent due to the expansion of one set of rods shall be equal to the rise due to the expansion of the other, or that the distance between A and F shall be the same at all temperatures, is

$$L'\alpha t = L''\beta t,$$

or

$$L'/L'' = \beta/\alpha.$$

Thus, if the total lengths of the rods of the two metals are to one another inversely as the coefficients of linear expansion of the materials, the length of the arrangement will be the same at all temperatures. If, then, the point A is the point of attachment of a pendulum, and the bob is attached to F, this arrangement will form a compensated pendulum. Since the coefficient of linear expansion of brass is $.187 \times 10^{-4}$, and that of steel $.11 \times 10^{-4}$, we have—

$$\frac{\textit{Total length of Steel Rods}}{\textit{Total length of Brass Rods}} = \frac{.187}{.11}$$
$$= 1.7.$$

A pendulum built up in this way is shown in Fig. 149, and is called a gridiron pendulum. The shaded rods are steel as well as the spring *b* and the middle rod *i*, which passes freely through holes in the two lower cross-pieces.

FIG. 149.
(From Ganot's "Physics.")

In the case of chronometers, the diameter of the balance-wheel increases with rise in temperature, and so its moment of inertia increases, causing the balance to oscillate more slowly. A more important effect is, however, caused by the influence of temperature on the elasticity of the hair-spring. As the temperature rises, the elasticity of the hair-spring decreases, and so the period of the balance-wheel decreases on this account also. In order to compensate for this change in the hair-spring, the moment of inertia of the balance-wheel ought to *decrease* as the temperature increases. This is brought about by making the rim of the balance-wheel of two strips of metal as shown in Fig. 150, the outer strip being composed of a metal with a higher coefficient of expansion than that composing the inner strip. The result is that, when the temperature rises, the outer strip expanding more than the inner, causes the semicircles AB and CD to become more curved, so that the ends B and D being fixed to the spokes of the wheel, the ends A and C move in towards the axis E, about which the wheel oscillates. In this way the moment of inertia of the wheel decreases as the temperature increases, and, by suitably adjusting the dimensions of the wheel, the shortening of the period on this account can be made exactly to compensate for the lengthening on account of the weakening of the hair-spring.

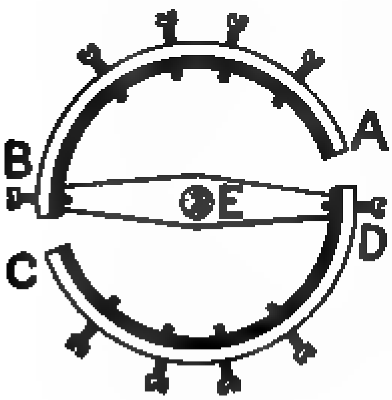


FIG. 150.

187. Cubical Expansion of Solids. — If we had a cube of a solid, of which the coefficient of linear expansion was a, each edge of which measured one centimetre at 0°, so that its volume at this temperature was 1 cc., and heated it to 1°, it would expand. The length of each edge would become $1 + a$, and hence the

volume would become $(1+a)^3$ or $1+3a+3a^2+a^3$. Now, as has been pointed out, the quantity a is small, so that a^2 and a^3 are so very small that we may, without making any appreciable error, neglect them, and call the volume of the cube at 1° $1+3a$. In the case of zinc, of which the coefficient of linear expansion is .0000292, the value of $3a^2$ is 2.5×10^{-9} or 0.0000000025, while the value of a^3 is 2.5×10^{-14}. Hence the volume at 1° when we include the terms in a^2 and a^3, is 1.0000876025, while omitting these terms the volume is 1.0000876.

The increase in the *volume* of unit volume of a body, when its temperature is raised 1°, is called the coefficient of cubical expansion of the body. Or, in other words, the increase in volume of a body, when its temperature is raised 1°, divided by its original volume, is called the coefficient of cubical expansion of the body. Hence we see that if a is the coefficient of linear expansion of a body, the coefficient of cubical expansion will be $3a$.

188. Expansion of Non-Isotropic Bodies.—In the preceding sections we have dealt with the expansion of isotropic bodies, in which the coefficient of linear expansion is the same in every direction. In the case of non-isotropic bodies, such as crystals, the coefficient of linear expansion is different in different directions. Thus, in the case of Iceland spar, the coefficient of linear expansion parallel to the axis (§ 401) is 0.2631×10^{-4}, while that perpendicular to the axis is 0.0544×10^{-4}. In the case of quartz, the coefficient of expansion parallel to the axis is 0.0797×10^{-4}, and that perpendicular to the axis is 0.1337×10^{-4}.

189. Coefficient of Expansion of Fluids.—In the case of fluids, we have only to deal with the coefficient of cubical expansion. Fluids, particularly gases, are so much more expansible than solids, that, as has been mentioned in § 184, we have, in their case, to make an addition to the definition of the coefficient of cubical expansion given for solids, and say that the coefficient of cubical expansion is the increase in volume of a given mass of the fluid when its temperature is raised one degree, divided by the volume of the same mass at 0° C. In calculating the increase in volume of, say, a volume V_1, of a fluid at a temperature t_1, when heated to t_2, we must include the terms in a^2 and a^3 (§ 184).

190. Expansion of Liquids—Apparent Expansion.—Suppose that a glass vessel A (Fig. 151) is filled with a liquid, say water, and that, when the whole is at a temperature of 0°, the volume of the water, *i.e.* of the glass vessel, is V_1; the surface of the water being at B. Suppose, now, it were possible to raise the temperature of the glass vessel to 1°, keeping the temperature of the contained water still at 0°. The vessel would expand, and its volume up to the mark B would become $V_1(1+\beta)$ where β is the coefficient of *cubical* expansion of the glass. Hence the water surface would fall to C, the volume of the bore between B and

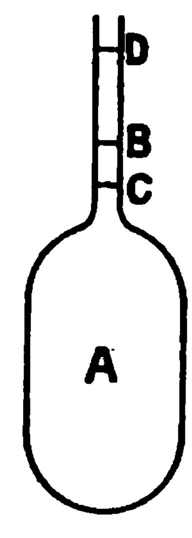


FIG. 151.

C being $V_1\beta$. If, now, the water was also heated to 1° it would expand, and its volume would become V_1 $(1+a)$, where a is the coefficient of expansion (cubical) of the water. Hence if the surface of the water now stood at D, the volume of the bore between C and D would be V_1a. Therefore the volume of the bore between B and D is $V_1(a-\beta)$. If we had not allowed for the expansion of the envelope, we should have taken the volume $V_1(a-\beta)$, *i.e.* the volume of the bore between the first and last positions of the surface of the water, as the expansion, and this volume divided by the volume at 0°, *i.e.* $V_1(a-\beta) \div V_1$, or $a-\beta$, as the coefficient of expansion of the liquid. The coefficient of expansion of a liquid obtained without considering the expansion of the containing vessel is called the *apparent* coefficient of expansion or dilatation, and it is equal to the difference between the true or real coefficient of expansion (a) of the liquid and the coefficient of cubical expansion (β) of the solid envelope.

The bulb shown in Fig. 151 is called a volume dilatometer, and is employed for determining the apparent expansion of liquids. The cubical expansion of the glass is obtained either by filling the bulb with a liquid of known absolute expansion, and making a series of measures, or by experiments on the linear expansion of a rod of the same glass. This known, the absolute expansion of the liquids can be calculated from the apparent expansion. The stem of the dilatometer is graduated, and the volume, up to the zero graduation, is determined by first weighing the dilatometer empty, and then when filled to the zero mark at 0° with mercury or water, and from the weight of mercury or water calculating the volume, using the known density of these liquids at 0°. The volume of the bore between two divisions is obtained in the same way by weighing a thread of mercury which, when in the stem, occupied a known number of divisions. The dilatometer is filled with the liquid, the expansion of which has to be measured, and the position of the surface in the stem noted when the dilatometer is placed in melting ice. The whole is then heated to a temperature $t°$ in a water bath, the temperature being measured by a thermometer placed in the water, and the position of the liquid surface again noted. The difference between the readings gives the expansion, and this quantity divided by the volume at 0° and by t gives the coefficient of apparent expansion.

Another form of dilatometer is the weight dilatometer. This resembles the volume dilatometer, except that there is only a single graduation on the stem, so that it resembles a specific-gravity bottle. The dilatometer is weighed empty; suppose its weight be w_0. It is then filled with the liquid and placed in melting ice. When the contents have reached a temperature of 0°, the quantity of liquid is adjusted by means of a capillary tube till the surface coincides with the mark on the stem. The dilatometer, full to the mark at 0°, is again weighed; let its weight be w_1. The dilatometer is now heated to a temperature of $t°$, and some

of the liquid withdrawn till the surface coincides with the mark, and the dilatometer weighed; let its weight be w_2. Now, neglecting for the present the expansion of the glass, a weight $w_2 - w_0$ of liquid at $t°$ occupies the same volume as a weight $w_1 - w_0$ at 0°. Hence, if V is the volume of the dilatometer, the density of the liquid at 0° is $\frac{w_1 - w_0}{V}$, and the density at $t°$ is $\frac{w_2 - w_0}{V}$. Now, the volume of a given mass is inversely proportional to the density. Hence the volume of a gram of the liquid at 0° will be $\frac{V}{w_1 - w_0}$, and the volume of the same mass at $t°$ will be $\frac{V}{w_2 - w_0}$. Hence the increase in volume will be

$$\frac{V}{w_2 - w_0} - \frac{V}{w_1 - w_0} \text{ or } \frac{V(w_1 - w_2)}{(w_2 - w_0)(w_1 - w_0)},$$

and the coefficient of apparent expansion, which is the increase in volume for 1° divided by the volume of the same mass (a gram) at 0°, *i.e.* by $\frac{V}{w_1 - w_0}$, is

$$\frac{w_1 - w_2}{w_2 - w_0}.$$

Then, subtracting the coefficient of cubical expansion of the glass from the apparent expansion, the coefficient of cubical expansion of the liquid is obtained.

Probably the most accurate method of determining the coefficient of expansion of a liquid is that which has been described in § 164, and has been employed by Benoist. By determining the loss of weight of the quartz cube in the liquid at different temperatures, the mass of a known volume of this liquid, the volume of the cube being known at these temperatures, is obtained. Hence the density at the different temperatures, and therefore, by a calculation similar to that given above, the coefficient of absolute expansion of the liquid can be calculated.

191. Direct Determination of the Absolute Expansion of Liquids.—A direct method of determining the absolute coefficient of expansion of mercury was devised by Dulong and Petit. The principle of this method is to use a U-tube filled with mercury, one limb being kept at 0°, and the other at 100°. If Δ_0 is the density of mercury at 0°, and Δ_{100} that at 100°, then if h_0 and h_{100} are the heights of the mercury in the two limbs above the horizontal part of the U, we get, by the principle of balancing columns (§ 145), $h_0\Delta_0 = h_{100}\Delta_{100}$.

Now the volume of unit mass of mercury at 0° is $1/\Delta_0$, and at 100° $1/\Delta_{100}$, so that the increase in volume is $\frac{\Delta_0 - \Delta_{100}}{\Delta_0 \Delta_{100}}$. The increase

divided by the volume at 0° ($1/\Delta_0$), and by the temperature range, gives the coefficient of expansion γ. Hence

$$\gamma = \frac{1}{100}\frac{\Delta_0 - \Delta_{100}}{\Delta_{100}} = \frac{1}{100}\left(\frac{\Delta_0}{\Delta_{100}} - 1\right).$$

$$\therefore \qquad \gamma = \frac{1}{100}\left(\frac{h_{100}}{h_0} - 1\right) \text{ or } \frac{1}{100}\frac{h_{100} - h_0}{h_0}.$$

The form of apparatus as used by Dulong and Petit was subject to numerous errors, but Regnault introduced some modifications, so that most of these were avoided, and his form of the apparatus only will be described.

Two upright iron tubes AA′, BB′ (Fig. 152), are connected together near the top by a horizontal tube C. At the bottom two horizontal branches A′D and B′F are connected to two vertical glass tubes DG and FH, these tubes being joined together at the top by a T-piece, the third arm of the T being connected to a glass globe K containing compressed air. The tubes AA′ and BB′ are surrounded by cylindrical vessels, which are filled with water or ice. The water in these cylinders is kept well stirred, and the temperature in the one which is heated is measured by an air-thermometer (§ 198) T, and that of the other by a series of mercurial thermometers t_1, t_2, t_3.

FIG. 152.

The upper horizontal tube AB is pierced at C with a small hole, so that the upper surfaces of the mercury in the vertical tubes are at the same level as C.

The pressures at A and B being equal, and those at G and H also equal, it follows that a column of mercury equal in height to the vertical distance (h) between H and G represents the difference in the pressures due to two columns of mercury, each of height AA′ or BB′, one hot and the other cold. Hence, if the temperatures of the cold column and of the columns DG and FH are the same, the effective height of the cold column is BB′ less the height GH, while the height of the hot column

is BB′ or H. Substituting these values in the expression for γ obtained above, we get

$$\gamma=\frac{1}{t}\left(\frac{h}{H-h}\right).$$

Where t is the temperature of the hot column, and the temperature of the cold column is taken to be zero. The mean value obtained for the coefficient of cubical expansion of mercury between 0° and 100° is 0.0001819.

192. Density of Water at different Temperatures—Point of Maximum Density.—The expansion of water with rise of temperature is anomalous, since this substance has a maximum density at a temperature of 4° C. If water is cooled below 4° it *expands*, and its density decreases, as shown in the table on p. 175.

This property possessed by water has an important bearing in nature, for otherwise all deep lakes in temperate zones would become frozen into a solid block of ice, and only the upper surface would thaw in the summer. As it is, in winter the surface water becomes cooled by radiation, &c., and as it cools, and its temperature falls, it becomes denser and sinks, convection currents being set up. This convection goes on till the temperature of the whole mass of water has fallen to 4° C. On the surface water becoming further cooled it expands, and its density becomes less than that of the water beneath. Hence the colder water remains on the top, and convection currents are not set up. Water being a very bad conductor of heat (§ 240), it takes a long time for the deeper layers of water to part with their heat, and so, even in the hardest winters, the ice in temperate zones is seldom very thick, and the water at the bottom of deep lakes is never colder than 4°.

The experiments made by Despretz in order to determine the temperature of maximum density illustrate very clearly the changes in temperature which take place in a mass of water, such as a lake. The apparatus, which is a modification of one devised by Hope, consists of a tall metal cylinder fitted with a lid, through the side of which are inserted four thermometers (Fig. 153). The cylinder was filled with water at a temperature of about 10° C., and was then placed in a cold room, of which the temperature was about 0°. The temperatures of the water at different depths, as indicated by the thermometers, were noted at frequent intervals, and the results plotted on curve paper, the ordinates being the temperature, and the abscissæ the times. In this way the curves shown in Fig. 154 were obtained. As the water near the sides lost its heat by radiation it became cooled and sank. Thus thermometer a sank the most rapidly, till it indicated a temperature of 4°. Next thermometers b, c, and d arrive at the same temperature in succession. The whole mass being at

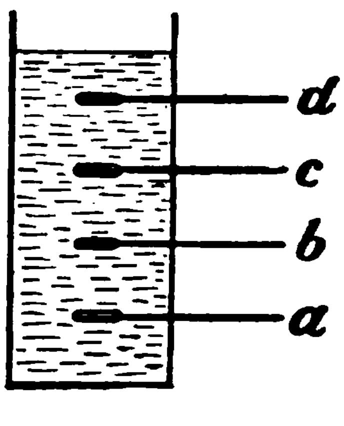


FIG. 153.

the point of maximum density, when the temperature fell any more, the colder water was lighter, and hence the upper thermometer (*d*) began to fall first, the others following in the order *c*, *b*, *a*. These changes are

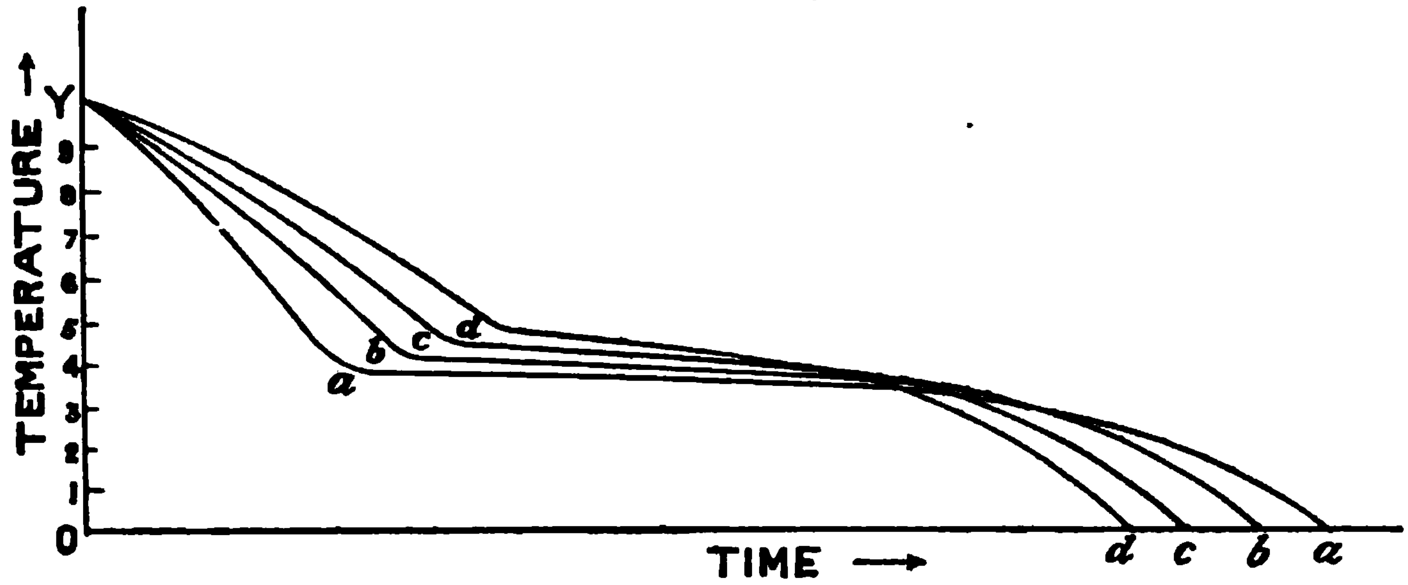


FIG. 154.

well indicated by the curves. The reason that the horizontal parts of all the curves are not coincident, is that disturbing currents are produced by the manner in which the water is cooled, namely, from the sides of the vessel.

193. Expansion of Gases—Expansion at Constant Pressure and at Constant Volume.—In the case of gases, owing to their compressibility, we have to take account of the pressure to which the gas is subject when its temperature is raised in order to get its thermal expansion. In determining the coefficient of expansion, we may either keep the pressure constant and measure the increase in volume when the temperature is raised, and thus obtain the coefficient of expansion at constant pressure, or we may vary the pressure so as to keep the volume constant, and from the amount by which the pressure has to be changed calculate the coefficient of expansion at constant volume. In either case, since the expansion of gases with temperature is considerable, we must refer the increase in volume or pressure, as the case may be, to the volume or pressure at 0° (§ 184).

194. Expansion of a Gas at Constant Pressure.—The first experiments on the expansion of a gas at constant pressure, having any pretensions whatever to accuracy, were made by Gay-Lussac, who measured the expansion of the gas contained in a glass bulb by the motion of a small mercury index. Owing, however, to the gases not being quite free from moisture, and to the fact that a mercury index does not completely enclose the gas, a film of gas existing between the mercury and the walls of the tube, the results obtained were not very accurate.

Regnault, having taken up the question, devised the form of apparatus shown in Fig. 155. The gas is contained in a glass bulb A, which is

connected by means of a very fine-bore tube with the graduated tube B. The lower end of this tube is connected with an upright tube C, which is open at the top. A tap D serves to draw off the mercury from the tube when required. The air is exhausted from the globe through a three-way cock E, the globe being heated to a temperature of at least 100°. Air or other gas is then admitted after passing through a series of drying tubes. The process of exhaustion and filling with dry gas is repeated a number of times, in order to remove from the inside of the globe the layer of moisture which is always condensed on a glass surface.

The globe having finally been filled with dry gas, the mercury is adjusted so that the surface is near the top of the graduations in B. The tap E is then turned in order to cut off the connection between the globe and the drying-tubes, the bulb is surrounded by melting ice or water at a known temperature, and the positions of the mercury surfaces in B and C noted. The pressure to which the gas is subjected is equal to the barometric height *plus* or *minus* the difference in level of the mercury in the tubes B and C.

FIG. 155.

The globe is then heated to a known temperature, say 100°, and the mercury run out of the tap D till the difference in level of the mercury in the two tubes, and hence the pressure acting on the gas, is the same as before. The volume of the tube B between the various graduations having been previously obtained, the amount of the increase in volume of the gas can be calculated. The water bath surrounding B and C serves to keep the mercury and the gas in B at a constant temperature, which is read off on the thermometer T.

195. Measurement of the Expansion of a Gas at Constant Volume.—As in the case of the coefficient of expansion at constant pressure, the first really accurate measurements of the coefficient of expansion of gases at constant volume were made by Regnault. One of the forms of apparatus which he employed for this purpose is shown in Fig. 156.

The gas is contained in a glass or porcelain globe A, which is connected by a fine tube with an upright tube B. A side tube and a three-way tap F allow of the globe being filled with dry gas, as in the constant pressure apparatus. At the upper part of the tube B an index E (shown on a larger scale at G), consisting of a curved piece of black glass, is attached

to the inside, and the mercury is always brought so that the surface just touches the point of this piece of glass; the pressure being adjusted by pouring mercury into the open tube C, or removing mercury through the stop-cock D. The difference in level, which is measured by a cathetometer, between the surface of the mercury in B and C, added to the barometric height, gives the pressure to which the gas is subjected.

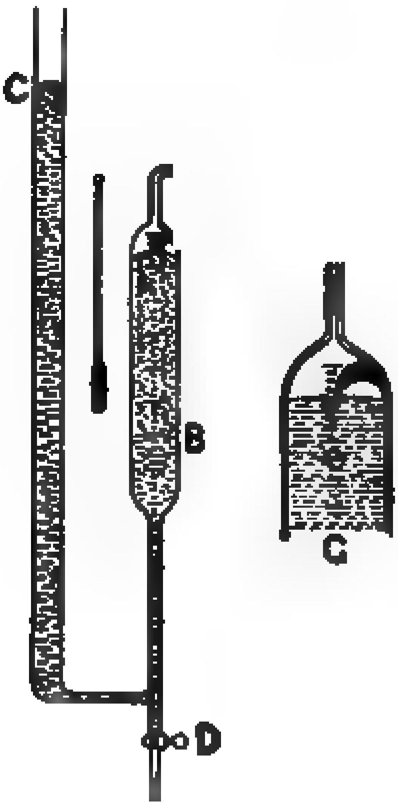


FIG. 156.

If the globe did not expand when heated, the volume of the gas would be the same at each temperature; owing, however, to the expansion of the globe a correction has to be applied, which, on account of the small coefficient of expansion of glass or porcelain compared with that of a gas, is not considerable.

196. Effect of Change of Pressure on the Coefficients of Expansion of Gases.—Regnault made a series of experiments on the coefficient of expansion at different pressures on air and some other gases, and the following table gives some of his results:—

COEFFICIENT OF EXPANSION OF GASES.

Gas.	Constant Pressure.		Constant Volume.	
	Pressure at 0° in cm. of Mercury.	Coefficient of Expansion.	Pressure at 0° in cm. of Mercury.	Coefficient of Expansion.
Air	76 cm.	.003671	11 cm.	.003648
"	257	.003695	76	.003665
"	...	...	200	.003690
"	...	...	2000	.003887
Nitrogen	...	...	76	.003668
Hydrogen	76	.0036613	76	.003669
"	254	.0036616	...	...
Carbon dioxide	76	.003710	76	.003686
" "	252	.003845	200	.003752
" "	...	...	800	.004252

An examination of the above table will show that in the case of air and carbon dioxide the coefficient of expansion increases as the pressure to which the gas is subjected increases. This increase of the coefficient of expansion being much more noticeable in the case of carbon dioxide, a gas which deviates considerably from Boyle's law (§ 130), than in the case of air; hydrogen, again, being an exception, in that the coefficient of expansion is constant, at any rate up to a pressure of 254 cm. of mercury. In most cases the coefficient of expansion at constant pressure is greater than that at constant volume. The most important characteristic, however, is that the coefficients of thermal expansion for the different gases are almost exactly the same, the differences between different gases being less, the lower the pressure to which the gases are subjected. We are therefore led to the conclusion that in the case of perfect gases (*i.e.* ones which obey Boyle's law exactly) the coefficients of thermal expansion would be constant, and have a value about 0.00366.

197. Charles's Law—Absolute Zero.—The law that all gases have the same coefficient of thermal expansion was first enunciated by Charles. It is, however, sometimes known as Gay-Lussac's law, since he was the first to make experiments to test the accuracy of the law. From what has been said in the previous section, it will be seen that gases obey Charles's law with about the same accuracy with which they obey Boyle's law.

If we consider a mass of a perfect gas of which the volume is v_0 and the pressure p_0, at a temperature of 0°C., and imagine the volume kept constant while its temperature is lowered to $-t°$, the pressure p will, by Charles's law, be given by

$$p=p_0(1-at),$$

where a is the coefficient of expansion. If the cooling is continued to a temperature of $\left(-\frac{1}{a}\right)°$, then

$$p=p_0(1-1)=0,$$

i.e. at this temperature the gas would exert no pressure on the walls of the containing vessel. According to the kinetic theory of gases, this can only occur when the velocity of translation of the molecules is zero. If, as seems probable, the motion of the atoms in the molecules, or perhaps it is better to say the vibratory motion of the molecules, increases and decreases *pari passu* with the motion of translation, it follows that at a temperature of $-\left(\frac{1}{a}\right)°$ C. the molecules will have completely lost all their motion. Heat consisting, as we shall see later, of the motion of the molecules, when no such motion exists the body must be devoid of heat. Since it is impossible to imagine a body colder than one which is devoid of all heat, *i.e.* one at a temperature of $-\left(\frac{1}{a}\right)°$ C., this temperature is called the *absolute zero.*

Taking a as .0036625 (the mean value for hydrogen between 0° and 100°), the absolute zero will be $-\frac{1}{.0036625}$, or $-273°.0$ C. Although it is impossible actually to cool a body down to the absolute zero, it is interesting to note that temperatures as low as $-250°$ C. have been obtained by allowing liquid hydrogen to boil at reduced pressure. The true value of such low temperatures is, however, difficult to estimate, since it is hardly safe to say that any property of matter which we may use to measure temperature will, at such low temperatures, change with temperature according to the same law as is found to hold at temperatures near 0° and 100° C.

In order to convert temperatures referred to 0° C. to the corresponding temperatures referred to the absolute zero, we have to add 273. Thus if T and t represent the temperature reckoned from the absolute zero and the temperature of melting ice respectively,

$$T = t + 273.$$

According to Charles's law,

$$p = p_0(1 + at),$$

and

$$v = v_0(1 + at).$$

Hence, substituting for a its value $1/273$, and reckoning the temperature from the absolute zero, the above equations become

$$p = p_0\left(1 + \frac{1}{273}(T - 273)\right)$$

$$= p_0 T/273,$$

and

$$v = v_0 T/273.$$

At any other temperature T', if the pressure when the volume is constant is p' and the volume when the pressure is constant is v', we have

$$p' = p_0 T'/273$$

and

$$v' = v_0 T'/273.$$

Hence

$$\frac{p}{p'} = \frac{T}{T'}, \text{ and } \frac{v}{v'} = \frac{T}{T'},$$

or the pressure of a gas at constant volume varies directly as the absolute temperature, and the volume, at constant pressure, also varies directly as the absolute temperature.

Suppose that the conditions of a certain mass of gas, as far as pressure, volume, and temperature are concerned, are indicated by the symbols p, v, t respectively, while when the temperature of the same mass of gas is reduced to 0°, the pressure being p_0, the volume is v'. Then if the temperature is maintained constant we have, by Boyle's law,

$$p_0 v' = pv,$$

or

$$v' = \frac{pv}{p_0}.$$

Now, keeping the pressure p_0 constant, let the temperature of the gas

be reduced to 0°, and let v_0 be the volume under these conditions. By Charles's law we have

$$v' = v_0(1 + at),$$

where a is the coefficient of expansion of the gas. Hence, equating the two values of v', we get

$$pv = p_0v_0(1 + at).$$

Taking the value of a as .003663, or $\frac{1}{273}$, and writing T for the temperature measured from the absolute zero, *i.e.* $-273°$ C., we get

$$pv = p_0v_0\left(1 + \frac{t}{273}\right) = \frac{p_0v_0}{273}(273 + t)$$
$$= \frac{p_0v_0}{273} . T.$$

For a given mass of gas the quantity p_0v_0 is a constant, hence we may write the above equation—

$$pv = RT,$$

where R is a constant, depending only on the nature and quantity of the gas.

198. The Gas Thermometer.—Since the standard thermometric substance employed for all accurate measurements of temperature is either hydrogen or nitrogen, the problem of comparing the readings of the ordinary liquid-in-glass thermometers, such as are actually used to note the temperature, with the gas thermometer, and hence deducing the corrections to be applied to the readings to reduce them to the gas scale, is of very considerable importance. There are several forms of so-called

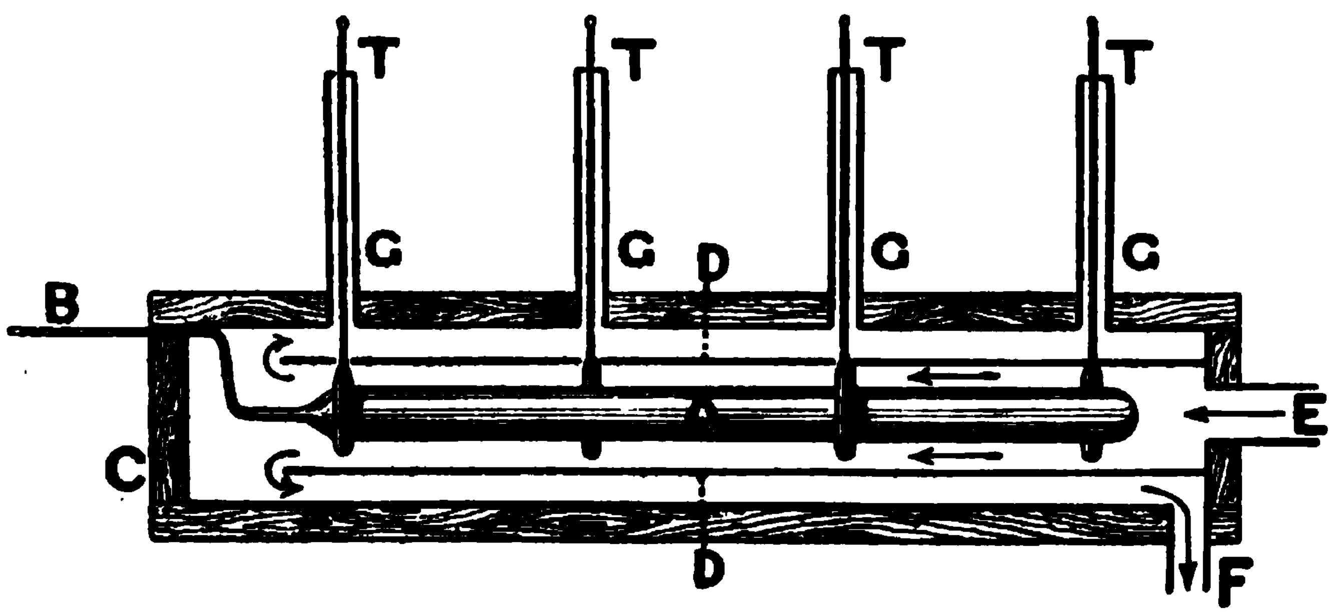


FIG. 157.

air thermometers, which are all more or less modifications of the instruments used by Regnault, and we shall content ourselves with describing the form employed at the Bureau International des Poids et Mesures at Paris.

The instrument consists of two distinct parts, the bulb, containing the gas (hydrogen), and the manometer, used to measure the pressure to

which the gas is exposed at the different temperatures. A section of the arrangement employed for heating the bulb A is shown in Fig. 157. The bulb A, which is made of platinum iridium, and has a capacity of about a litre, is connected with the manometer (shown in Fig. 158) by a fine metal tube B, about a metre long, and having a bore of 0.07 cm. For the comparisons at comparatively low temperatures the bulb A and the thermometers T, which are to be compared with the gas thermometer, are placed side by side in a long water-bath, which is kept well stirred. For the higher temperatures the arrangement shown in Fig. 157 is employed. Steam, or the vapour of some other liquid, enters the apparatus by the tube E, passes up alongside the bulb A and the bulbs of the thermometers T, and then at the end passes to the outside of the metal screen DD and back along the outside, finally escaping by the tube F. The arrangement resembles that used for determining the upper fixed point of an ordinary mercurial thermometer (see Fig. 144).

In Regnault's form of the constant-volume air thermometer, the manometer employed only measures the excess or defect of the pressure to which the gas is exposed over the ordinary atmospheric pressure, so that to obtain the actual pressure the barometric height has also to be determined. In the Bureau instrument the manometer and barometer are combined in a single instrument, so that the height of a single column of mercury only has to be measured, and thus the chances of error are reduced. The tube B, coming from the bulb, is attached to a steel block A (Fig. 158), which is clamped air-tight on the end of a glass tube C. The lower end of this glass tube is cemented into a steel block D, to which is also cemented a second glass

Fig. 158.

tube E. These two glass tubes communicate with each other through a channel in the steel block, as well as with a tap and flexible steel tube K. The block D is attached to an upright metal pillar P, which also carries a movable cradle Q, the position of which can be adjusted by the screw S. The cradle Q carries the upper end of a barometer tube HG, the lower end of which dips in the mercury contained in the tube E. The lower surface of the steel plug A is made plain, except for a fine metal point, shown on a larger scale at N, which serves as a fixed mark to which the surface of the mercury in the tube C is always brought back. The height of the reservoir L is altered, roughly by sliding the cradle R up and down by hand, and finally by means of the screw M, till this adjustment is complete at each temperature.

When the surface of the mercury at J is exactly in contact with the steel point, the excess of the pressure within the bulb above the atmospheric pressure is equal to the weight of a column of mercury of height OJ. The atmospheric pressure is equal to the weight of a column of mercury, of height IO. Hence the pressure acting on the gas in the bulb is equal to the weight of a column of mercury of height IO + OJ or IJ, and the measurement of the vertical distance between the two mercury surfaces I and J suffices to give the pressure. The measurement of this height is effected by means of a cathetometer, which is carried on a pillar fixed alongside the instrument, the measurement being facilitated by the fact that the two surfaces I and J are placed vertically one over the other. The temperature of the mercury column is measured by a series of thermometers attached to the upright P.

The readings obtained have to be corrected to allow for the expansion of the bulb on account of (1) rise of temperature and (2) the increase of the pressure of the gas inside. Allowance has also to be made for the decrease in volume, as the pressure is increased, of the air contained in the tube BB and the space between the mercury meniscus J and the lower surface of the steel block A. The coefficient of cubical expansion of the platinum-iridium bulb was determined by measuring, directly on the comparator, its coefficient of linear expansion.

CHAPTER II

CALORIMETRY

199. Quantity of Heat.—In order to measure the quantity of heat which a body loses or gains when its temperature changes, or when its physical state changes, we generally use as the unit that quantity of heat which, acting on a given mass of some chosen substance, alters its temperature by some fixed amount. The substance employed almost exclusively for this purpose is water. Thus the unit of heat might be defined as the heat necessary to raise the temperature of one gram of water through one degree Centigrade. This definition, however, will only be complete if a gram of water requires the same quantity of heat to raise its temperature one degree, whatever the temperature at which we start; that is, if it required the same quantity of heat to raise a gram of water from 0° to 1°, as from 15° to 16°, or from 90° to 91°. Since it has been found that the quantity of heat required at different temperatures is different, it is necessary to specify between what two temperatures the water has to be taken, and there are a number of thermal units in use differing from one another in the temperature at which the water is taken. The chief of these are as follows :—

(1) The heat required to raise 1 grm. of water from 0° C. to 1° C.

(2) The heat required to raise 1 grm. of water from 3°.5 C. to 4°.5 C.

(3) The heat required to raise 1 grm. of water from 14°.5 C. to 15°.5 C.

Each of the above units has, at some time or other, been called a *calorie*, and so in accurate work it is necessary to say at what temperature the calorie is taken.

A unit of heat largely used in England in engineering is the heat required to raise 1 lb. of water through 1° F. As this unit is only used for comparatively rough measurements, the question as to the temperature at which the water is taken does not come in.

For theoretical purposes (and practical also, now that electrical measurements play such an important part in engineering), it is convenient to measure heat in terms of the units of work or energy, since, as will be seen later (§ 250), heat and energy are convertible, and it has been proposed to adopt as the practical unit of heat 4.2×10^7 ergs—the reason for the adoption of this number will be seen later (§ 251)—and to call this unit a joule.

200. Specific Heat.—If 100 grams of water at 100° is mixed with 100 grams of water at 0°, the temperature of the mixture is found to be

very nearly 50°. But if 100 grams of copper at 100° is placed in 100 grams of water at 0°, the final temperature of the copper and water is not 50°, but about 9°.1. The heat given out by the copper in cooling from 100° to 9°.1, *i.e.* through 90°.9, has only been able to heat an equal mass of water through 9°.1. It is therefore evident that a given mass of copper requires much less heat to raise its temperature 1° than does an equal mass of water. The quantity of heat necessary to raise the temperature of 1 gram of a substance through 1° C., at any given temperature, is called the *specific heat* of the substance at that temperature. The specific heat of water (at the temperature at which the calorie is defined) is therefore unity, and that of copper .1.

The following table gives the specific heat of some substances in terms of water at 15° C. The second column gives the mean temperature at which the specific heat of the substance was measured.

SPECIFIC HEAT.

Substance.	Temperature.	Specific Heat.
Water	5°	1.0041
"	10°	1.0019
"	15°	1.0000
"	20°	.9987
"	27°	.9967
Ice	− 10°	.502
Paraffin (wax)	10°	.694
Copper	50°	.092
Zinc	50°	.093
Iron	15°	.109
Platinum	50°	.032
Mercury	20°	.0331
Petroleum	40°	.51

201. The Measurement of the Specific Heat of Solids.—The most usual method of determining the specific heat of a solid is called the method of mixtures, and consists in heating a given mass of the solid to a known temperature, and then immersing it in a vessel containing a known mass of water, the initial and final temperatures of which are noted. If M is the mass of the body, W that of the water, t_1 the initial temperature of the body, t_2 the initial temperature of the water, and t_3 the final temperature of the body and the water, we have, if we suppose for the moment that the vessel does not take up any heat, that the heat gained by the water is $W(t_3 - t_2)$. The heat lost by the body is $M(t_1 - t_3)\,s$, where s is the specific heat of the body. Equating these two quantities of heat, we get

$$s = \frac{W(t_3 - t_2)}{M(t_1 - t_3)}.$$

Since the temperature of the vessel in which the water is contained

(called the Calorimeter) is raised from t_2 to t_3, some of the heat given out by the body will have been used for this purpose, and the above result must be corrected on this account. If w is the mass of the calorimeter, and σ the specific heat of the material of which it is composed, the heat necessary to raise its temperature from t_2 to t_3 is $w(t_3 - t_2)\sigma$. The product $w\sigma$, which represents the quantity of water which would require the same quantity of heat to raise its temperature 1° as does the calorimeter, is called the *water equivalent* or *water value* of the calorimeter. The heat gained by the water and calorimeter is $W(t_3 - t_2) + w(t_3 - t_2)\sigma$, and hence

$$s = \frac{(W + w\sigma)(t_3 - t_2)}{M(t_1 - t_3)}.$$

In forming the above expressions, we have supposed that *all* the heat given out by the hot body is received by the calorimeter and its contents. Since the hot body has to be moved from the enclosure in which it was heated to the calorimeter, special precautions have, as we shall see later on, to be taken to prevent loss of heat during transit. Again, although the calorimeter may be at the same temperature as its surroundings at one temperature, say the initial temperature, yet when the hot body is placed in, its temperature will be higher than that of its surroundings, and hence it will lose heat by conduction and radiation (§ 241). In order to reduce such loss of heat to a minimum, the calorimeter is supported on small, badly conducting feet, or suspended by threads, so that it shall not gain or lose heat by conduction through the supports. The loss or gain of heat by radiation is reduced to a minimum by having the outside of the calorimeter polished, since polished metal is a bad radiator (§ 246).

Rumford first proposed to eliminate the effects of radiation by making a preliminary experiment to determine approximately the rise in temperature of the calorimeter, and then, in the final experiment, cooling the calorimeter before the introduction of the hot body to a temperature below that of the surrounding bodies by an amount equal to half the rise. By this device, during the first part of the time between the introduction of the hot body and the reading of the final temperature, the calorimeter would receive heat by radiation, and during the second part it would lose heat. As, however, the temperature rises most rapidly at first, this correction is not perfect, and for very accurate observations the following method, adopted by Regnault, in which the loss or gain by radiation is directly measured, is used.

The temperature of the calorimeter and its contents is read at short intervals (τ), say every 10 seconds, after the introduction of the hot body until the maximum reading has been obtained, and the temperature begins to fall. The calorimeter is then left, and the fall of temperature in two or three minutes determined, and from this the fall in the

interval τ (10 seconds) calculated. This gives the rate of loss of heat at the highest temperature. The calorimeter is allowed to cool, and the rate of loss of heat in a time τ determined at a number of temperatures between the maximum and the original temperature. A curve PQ (Fig. 159) is then plotted, in which the abscissæ are the temperatures of the calorimeter and the ordinates are the fall in temperature at the different temperatures during the interval τ. The point P, where this curve cuts the axis of X, of course corresponds to the calorimeter, being at the same temperature as its surroundings, so that it neither gains nor loses heat by conduction or radiation. The readings of the temperature of the calorimeter while the temperature was rising are plotted on square paper, the abscissæ being the times from the instant when the hot body was introduced into the calorimeter. In Fig. 160 DBC is such a curve. An ordinate NR is then drawn corresponding to a time $\tau/2$ from the start, and the corresponding temperature T read from the curve. Next the fall of temperature (t_1) during the time τ, when the temperature of the calorimeter is T, is read from the curve (Fig. 159). This quantity t_1 is added on to the ordinate MD to give a new point D′. This point represents what would have been the temperature, if there had been no loss of heat, at the end of the time τ, for MD represents the actual temperature, and DD′ represents the loss which has taken place up to then. In the same way the fall of temperature t_2 corresponding to the temperature at the middle of the next interval is read off, and the sum of the two quantities t_1, t_2 is added to the ordinate EP, since the total fall of temperature due to radiation and conduction up to the time P is the loss during the interval OM, together with that during the interval MP. Proceeding in this way a

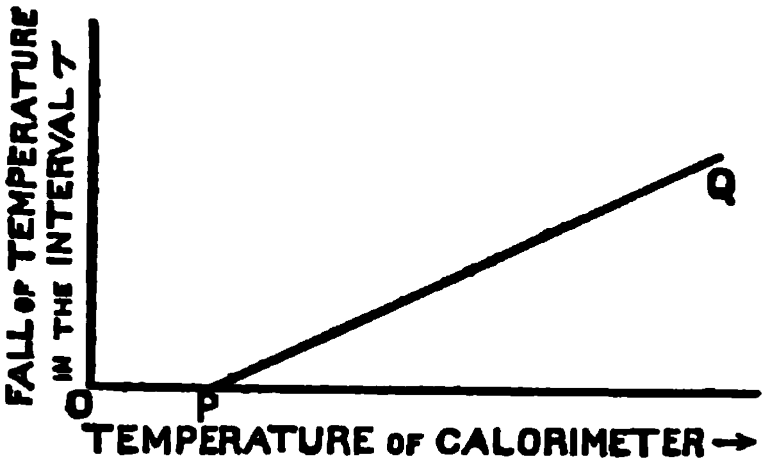


FIG. 159.

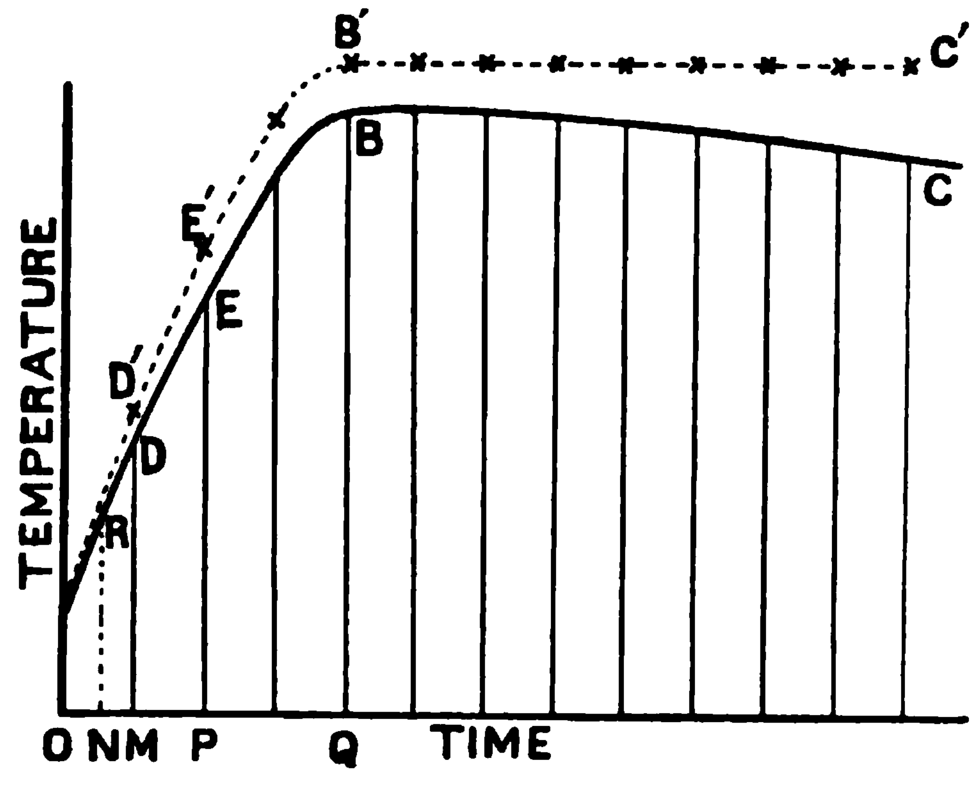


FIG. 160.

series of points are obtained, such as are indicated on the dotted curve D′E′B′C′. This curve rises to the point B′, and after that is horizontal. The reason why it remains horizontal is that we are allowing for the fall of temperature due to radiation, and hence the dotted curve represents the temperature of the calorimeter supposing there were no loss by radiation, under which circumstances the temperature would remain constant as soon as the hot body and the calorimeter had reached the same temperature. The final temperature taken in the calculation of the specific heat is that corresponding to the horizontal part of the dotted curve, *i.e.* QB′.

In Fig. 161 a modified form of Regnault's calorimeter is shown. The substance of which the specific heat is to be measured is heated in the heater A. This heater is shown in section at (*a*), and is connected by the side-tube E with a boiler, so that steam enters at E and passes out through F to a condenser. The temperature to which the substance is heated is indicated by the thermometer T_1. The calorimeter C is suspended by means of three light strings inside a brightly polished metal vessel D, while this vessel is itself contained within a wooden box B. A delicate thermometer T_2, which is held in a clip attached to the box B, is used to give the temperature of the liquid in the calorimeter, while a stirrer S serves to mix the liquid and thus insure it all being at the same temperature. A screen K, which slides up and down in guides, serves to protect the calorimeter from radiation from the heater. When the substance has attained the temperature of the heater, the screen K is raised, the box B run on its guides under the heater, and the substance dropped down into the calorimeter, the small slide L being momentarily drawn out for this purpose. Directly the substance has been introduced, the calorimeter is withdrawn, and the screen K again lowered.

FIG. 161.

The consideration of calorimetric methods which depend on latent heat of vaporisation or fusion will be dealt with later (§§ 212, 215).

Favre and Silbermann used a calorimeter which was essentially a

very large mercurial thermometer. The bulb consisted of an iron sphere connected to a narrow glass graduated stem. Into this sphere one or two closed tubes made of glass or platinum, somewhat of the shape of test-tubes, protruded. The instrument was standardised by introducing into one of the tubes a known weight of hot water, and noting the fall of temperature of the water and the distance through which the mercury column in the stem travelled. Then, from the advance of the column when another body was introduced into one of the tubes, the quantity of heat it imparts to the mercury could be calculated.

202. The Measurement of the Specific Heat of Liquids.—The method of mixtures is applicable in the case of liquids ; either a solid of known specific heat being used, the calorimeter containing the liquid, or, if the liquid does not combine chemically with water, a known mass of the liquid, at a temperature higher or lower than that of the water is run into the calorimeter.

203. The Measurement of the Specific Heat of Gases.—When a body expands it drives back the atmospheric pressure and hence does work (§ 252), and, as we shall see later (§ 253), this work is done at the expense of some of the heat supplied to the body. Thus the specific heat, *i.e.* the heat required to raise the temperature of unit mass 1° C., of a body will be different according as it is allowed to expand, and hence do external work, or kept at constant volume by suitably altering the pressure. In the case of solids and liquids, the expansion is so small that the external work done and the heat necessary to do this work are negligable. The specific heats as determined are at constant pressure but would differ inappreciably from the specific heats at constant volume. In the case of gases, where the change of volume when they are heated at constant pressure is considerable, the amount of heat required to do the external work performed by the expanding gas amounts to a large fraction of the heat supplied. Hence there are two specific heats to be considered in the case of a gas—(1) The specific heat at constant pressure, which is the heat required to raise the temperature of unit mass of the gas through 1° when the pressure is kept constant. (2) The specific heat at constant volume, which is the heat required to raise the temperature of unit mass of the gas through 1° when the volume is kept constant.

204. The Measurement of the Specific Heat of a Gas under Constant Pressure.—Accurate measurements of the specific heats of gases under constant pressure have been made by Regnault. The gas to be experimented upon was stored under pressure in a large metal reservoir A (Fig. 162). From this reservoir the gas passes along a tube to a screw valve B, shown in section at C. A little way beyond the valve there is a partition across the tube, pierced with a small hole D. Between this partition and the valve a side-tube leads to a manometer E. As the

gas escapes from the reservoir the pressure becomes reduced, but by opening the valve the pressure to the left of the partition, as shown by the manometer, can be kept constant, and hence the gas made to flow

FIG. 162.

through the apparatus at a uniform rate. The gas next passes through a long spiral F (Fig. 163) immersed in an oil bath. Having thus acquired the temperature of the bath, the gas passes into the vessel G contained in the calorimeter H. By means of a series of spiral partitions, as shown at K (Fig. 162), the gas is obliged to go round and round, so that it becomes cooled down to the temperature of the water in the calorimeter before escaping.

FIG. 163.

The mass of the gas which passes through the apparatus is obtained by noting the pressure in the vessel before and after the experiment. The volume having been previously determined, this allowed the mass to be calculated.

Corrections have to be applied for the loss of heat by radiation, and for heat conducted to the calorimeter by the tube through which the gas enters.

The following table contains some of the values obtained by Regnault :—

SPECIFIC HEAT OF GASES AT CONSTANT PRESSURE.

Air	0.2374
Chlorine	0.1220
Carbon dioxide	0.2169
Hydrogen	3.4090
Nitrogen	0.2438
Oxygen	0.2175

205. Specific Heat of Gases at Constant Volume.—The direct determination of the specific heat of a gas at constant volume is rendered very difficult from the necessity of enclosing the gas in a vessel, the water value (§ 201) of which will be enormously greater than that of the enclosed gas. Direct determinations of this quantity have, however, been made by Joly, who employed for this purpose the steam calorimeter described in § 215.

As will be seen later, it is possible, by measuring the velocity of sound in a gas, to calculate the ratio of the specific heat at constant pressure to that at constant volume. Knowing the specific heat at constant pressure, we are then able to calculate that at constant volume.

206. Variation of Specific Heat with Change of Temperature, Density, and State.—Regnault examined the values of the specific heat at constant pressure of gases, at different temperatures, and found that while the specific heat of air is practically constant, that of carbon dioxide increases considerably as the temperature rises. It is probable that all gases which deviate from Boyle's and Charles's laws show an increase of specific heat with increase of temperature, but that a perfect gas would possess a constant specific heat.

In the case of water, Rowland and Bartoli and Stracciati find that the specific heat decreases from 0° to a temperature of 30° (Rowland) or 20° (Bartoli and Stracciati), and then increases. The recent measurements of Griffiths, however, seem to show that there is no minimum at any temperature below 30°.

The specific heat of most solids increases with increase of temperature. The most noteworthy cases of the increase of specific heat with temperature are the solids carbon, boron, and silicon. For the reasons given in the next section, a study of the specific heat of these three bodies is of particular interest, and was undertaken by Weber, who employed Bunsen's ice-calorimeter (§ 212) for temperatures up to 200°, and a water-calorimeter for higher temperatures. The temperature to which the body was raised was obtained by having a lump of platinum heated to the same temperature as the body, and placing this in a second calorimeter. Then from the rise in temperature produced by the

platinum, and from the specific heat of platinum, the initial temperature was calculated. The values he obtained for carbon in the form of diamond are shown by a curve in Fig. 164. It will be noticed that the specific heat of diamond is about three times as great at a temperature of 300° as it is at 0°, while at temperatures above 600° the specific heat remains

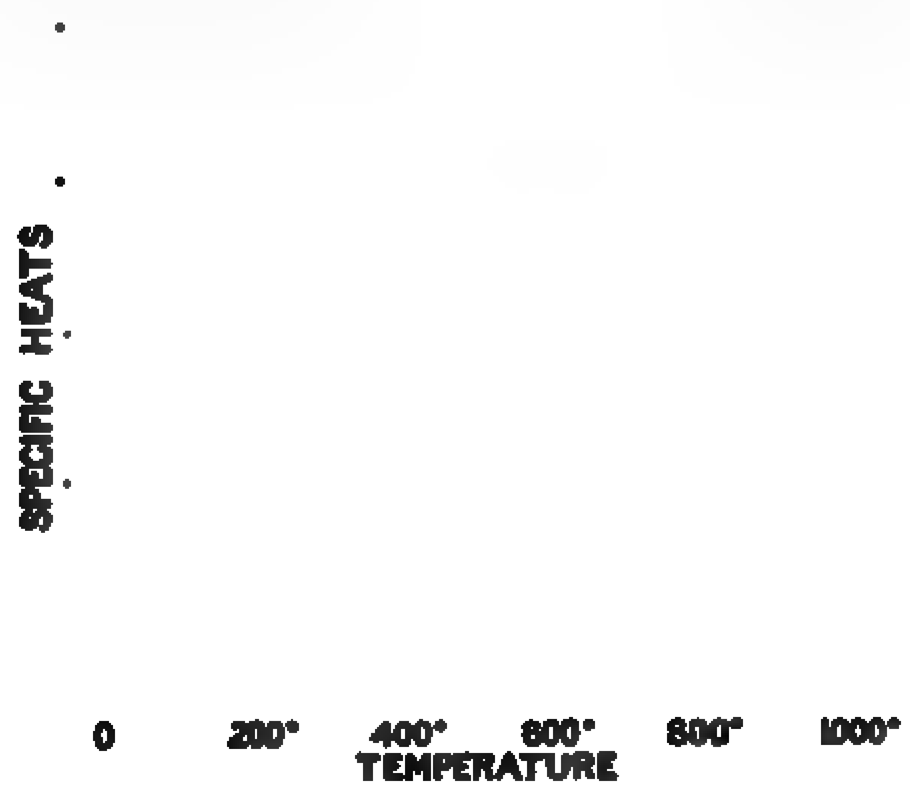


Fig. 164.

almost constant. Similar results were obtained in the case of boron and silicon, except that in the case of the latter substance the specific heat is almost constant at temperatures above 200°.

In the case of bodies which are capable of existing in more than one allotropic modification, marked differences in the specific heats of the various forms are often found. Thus in the case of calcium carbonate the specific heats of aragonite and Iceland spar are 0.2085, that of chalk 0.2148, and that of marble 0.2158. At ordinary temperatures the specific heat of carbon in the form of diamond is 0.1469, wood charcoal 0.2415, and graphite 0.2017. Weber found, however, that at high temperatures all forms of carbon have the same specific heat, the same probably being the case with other polymorphous bodies.

The specific heat of most bodies is different in the three states—solid, liquid, and gas. In general the specific heat in the solid and gaseous states are much smaller than in the liquid state. The following table gives some values of the specific heat for bodies in different states :—

CHANGE OF SPECIFIC HEAT WITH CHANGE OF STATE.

	Solid.	Liquid.	Gas.
Water	0.50	1.000	0.477
Mercury . . .	0.0314	0.0333	...
Tin	0.0562	0.0637	...
Lead	0.0314	0.0402	...
Alcohol . . .	...	0.5475	0.4534
Ether	...	0.5290	0.4797

207. Dulong and Petit's Law.—Dulong and Petit first enunciated the law that the product of the specific heat of an element in the solid state into the atomic weight is a constant. The product of the atomic weight into the specific heat of a gas is also approximately constant, but about half the value of the product in the case of solids. In the case of liquids the law does not apply at all. If n is the number of molecules in unit mass of a solid element, w the mass of each molecule, σ the heat required to raise the temperature of a single molecule 1°, and S the specific heat of the body, we have $S=\sigma n$ and $nw=1$. Hence $\sigma=wS$. Since, according to Dulong and Petit's law, wS is a constant, it follows that σ is also constant, or the heat required to raise the temperature of a molecule of all the elements, when in the solid state, through 1° is the same. The product of the specific heat into the atomic weight of an element is called the atomic heat, and the values of this quantity for some elements are given in the following table:—

ATOMIC HEATS.

	Atomic Weight.	Specific Heat.	Atomic Heat.
Hydrogen (gaseous)	1	3.409	3.4
Oxygen (gaseous)	16	0.218	3.5
Nitrogen (gaseous)	14	0.244	3.4
Iron	56	0.109	6.1
Copper	63	0.092	5.8
Zinc	65	0.093	6.1
Platinum	194	0.032	6.2
Arsenic	75	0.081	6.1
Selenium	79	0.084	6.6
Sodium	23	0.293	6.7
Potassium	39	0.170	6.6
Sulphur	32	0.163	5.2
Mercury (solid) . . .	200	0.031	6.2
Carbon	12	0.144	(1.7)
		(985°) 0.459	5.5
Boron	11	(27°) 0.238	(2.6)
		(233°) 0.366	4.0
Silicon	28	(57°) 0.183	(5.2)
		(232°) 0.203	5.7

It will be seen that the atomic heats of gases are about 3.4, and those of solids about 6.4. After what has been said in the previous section as to the change of specific heat with temperature and with the allotropic state of a body, the differences obtained are not surprising. The values of the atomic heats of carbon, boron, and silicon obtained by using the specific heat as measured at ordinary temperatures are very different from 6.4, and it was with a view to testing whether these abnormal values of the atomic heat would persist at all temperatures that Weber undertook his investigation into the specific heat of these bodies. It is to be noted that these three bodies are all very difficultly fusible, so that at ordinary temperatures they are a long way from their melting-point. The specific heat of most solids seems to become constant near a certain temperature, and hence it is only reasonable to employ the specific heat measured at such temperature for getting the atomic heat; and it is probable that, if this were done, Dulong and Petit's law would be more nearly true. In the case of carbon, boron, and silicon the table shows how very much better the atomic heats calculated from the specific heats at high temperatures agree with the other atomic heats, than do those calculated from the specific heats at low temperatures.

An extension of Dulong and Petit's law is due to Woestyn, who suggests that the atoms of the elements, even when combined with one another, preserve the same specific heat that they have in the uncombined state, so that the thermal capacity of the molecule of any compound is equal to the sum of the atomic heats of its constituent atoms. This law is not verified by experiment with any degree of completeness, and Neumann has limited the law to the statement that, for compounds belonging to the same general formula, and which are similarly constituted, the product of the molecular weight into the specific heat is constant; but that the value of the product varies from one series to another.

The following table gives the specific and molecular heats of some compounds :—

MOLECULAR HEATS.

Type RCl.	Molecular Weight.	Specific Heat.	Molecular Heat.
Sodium chloride ($NaCl$) . .	58.5	0.214	12.5
Potassium chloride (KCl) . .	74.5	0.173	12.9
Silver chloride ($AgCl$) . . .	143	0.091	13.0
Type RCl_2.			
Barium chloride ($BaCl_2$) . .	208	0.090	18.7
Strontium chloride ($SrCl_2$) . .	158	0.120	19.0
Lead chloride ($PbCl_2$) . . .	278	0.066	18.3
Type RSO_4.			
Barium sulphate ($BaSO_4$) . .	233	0.113	26.4
Strontium sulphate ($SrSO_4$ . .	184	0.143	26.3
Lead sulphate ($PbSO_4$) . . .	303	0.087	26.4

If we assume that the atom of sodium, potassium, or silver requires the same amount of heat to raise its temperature one degree, whether it is free or combined with chlorine, we can calculate the atomic heat of chlorine in the solid condition. Thus

$$12.8 - 6.4 = 6.4,$$

so that the atomic heat of solid chlorine is 6.4, and its specific heat is 6.4/35.5, or 0.18.

In the same way, from the salts of the type RCl_2, we have that the atomic heat of chlorine is

$$\tfrac{1}{2}(18.7 - 6.4) = 12.3/2 = 6.15.$$

Applying the same method of calculation to obtain the atomic heat of solid oxygen from the molecular heats of the salts of the type RSO_4, we get

$$\tfrac{1}{3}(26.4 - 6.4 \times 2) = 13.6/3 = 4.5.$$

In this case the atomic heat is distinctly below 6.4, and the mean value for solid oxygen obtained from oxides and salts is 4.1.

Assuming the value 4.1 for the atomic heat of oxygen, we can calculate the atomic heat of solid hydrogen. The specific heat of ice is 0.5, so that the molecular heat is $18 \times .5 = 9$. Hence the atomic heat of solid hydrogen is

$$\tfrac{1}{2}(9 - 4.1) = 2.5.$$

The numbers obtained in this way, depending as they do on so many hypotheses, are probably only approximately correct; they represent, however, an interesting application of Woestyn's extension of Dulong and Petit's law.

CHAPTER III

CHANGE OF STATE

208. Melting-Point.—One of the fixed points chosen for thermometry is the temperature at which ice melts when subjected to atmospheric pressure. As long as the ice is pure this temperature seems quite constant, and therefore is suitable for use as a fixed point. If a mixture of ice and water is at any temperature except 0° C., it will gradually change its physical state, some of the ice becoming fluid, if the temperature is above 0°, or some of the water solid if the temperature is below 0°. At a temperature of 0°, however, solid water, or ice, and liquid water can coexist without change. In the case of ice, the *melting-point*, which is the temperature of melting or the temperature at which water solidifies, is very well marked; there are, however, other bodies, such as glass, iron, &c., which, when heated, become gradually softer and softer as the temperature rises, passing through the conditions of a soft solid and a viscous fluid, so that they have no very well-marked melting-point, the solid passing into the liquid condition by insensible gradations.

The following table gives a series of melting-points, the temperature corresponding to the bodies at the upper end of the list being rather doubtful:

Fig. 165.

MELTING-POINTS.

	Deg. C.		Deg. C.
Iridium . .	2230	Zinc . . .	415
Platinum .	1800	Bismuth .	268
Copper . .	1096	Sulphur . .	115
Gold . .	1092	Paraffin .	52
Silver . .	985	Ice . . .	0
Aluminium	625	Mercury .	−39

209. Change in Volume during Fusion.—Most substances occupy a larger volume in the liquid than in the solid state, so that expansion takes place on solidification. There are, however, exceptions; some substances, such as ice, cast iron, and bismuth, expand on solidification. In the case of water, the changes in volume which take place between a temperature of − 20° C. and 50° are shown in Fig. 165.

The density of ice at 0° being 0.91674, and that of water at the same temperature 0.99987, the increase in volume of one gram of water when it solidifies is 0.0907 c.c. Water, when changing to ice, is capable of exerting an enormous force if its expansion is resisted. This expansive force is evident in the bursting of water-pipes, and the disintegration of rocks into the pores of which water has permeated.

210. Effects of Pressure on the Melting-Point.—In 1849 Prof. James Thomson showed[1] that it followed, from the mechanical theory of heat, that if a body expands on solidification, like water, then increasing the pressure will lower the freezing-point; while if the body contracts on solidification, like paraffin, then increasing the pressure will raise the freezing-point. He calculated that in the case of water, increasing the pressure by one atmosphere would lower the freezing-point by 0°.0075 C. Hence, under a pressure of 1000 atmospheres, water would not freeze above a temperature of −7°.5. In other words, if water remains liquid, which it must, unless it is able to expand as it passes into ice, at a temperature of −7°.5, then it must be subjected to a pressure of at least 1000 atmospheres, and it is clear how water in freezing is able to burst even thick steel shells.

The following table gives some of the results obtained by Tammann on the effect of pressure on the melting-point:—

CHANGE OF THE MELTING-POINT WITH PRESSURE.

Substance.	Pressure in kilograms weight per square centimetre.	Melting-Point. Deg. C.
Benzene . . .	0	5.3
"	500	19.0
"	1000	31.4
"	2000	54.8
"	3000	73.5
"	3500	81.4
Nitrobenzene . .	1	5.17
" . . .	500	16.2
" . . .	1000	27.4
" . . .	2000	46.4
" . . .	3000	65.4
" . . .	3500	74.3
Phosphorus . .	0	43.9
" . . .	500	57.8
" . . .	1000	71.5
" . . .	2000	97.4

It will be noticed that the change in the melting-point is very considerable. Thus at a pressure of 3500 kilograms per square centimetre

[1] See § 263.

(3387 atmospheres), the melting-point of benzene is above the ordinary boiling-point (81°) at a pressure of one atmosphere.

If a wire loop is passed round a block of ice, and a weight is attached to the bottom of the loop, it will be found that the wire gradually cuts through the block, but that the ice joins together again after the wire, so that the block remains whole. This phenomenon, which is referred to as regellation, is explained as follows. The whole block being at o°, the ice immediately under the wire is compressed, and has its melting-point lowered so that it can no longer remain solid at o°, and therefore it melts, letting the wire down, and the water flowing round the wire. The melting of the ice causes a lowering of temperature on account of the latent heat (§ 211) of fusion of the ice. The water, when it gets above the wire, is no longer compressed, and hence, as its temperature is below o°, it again freezes, joining together the severed portions of the ice above.

211. Latent Heat of Fusion.—When a vessel containing a mixture of ice and water at o° is heated, it is found, if the contents are well stirred, that the temperature remains at o° as long as any ice is left. Since heat is being supplied, and the temperature does not rise, it follows that heat must be required to convert ice at o° into water at the same temperature. This heat, which is employed not in changing the temperature of a body, but in changing its state, is called *latent heat*. In the same way, to convert water at o° into ice at o°, heat has to be abstracted. The quantity of heat required to melt unit mass of a solid, or the quantity of heat which must be removed to convert unit mass of a liquid into a solid, in both cases without changing the temperature, is called the *latent heat of fusion* of the body.

We have seen that, according to the molecular theory of the constitution of matter, the molecules in a solid are more closely held together than in a liquid, so that part at any rate of the latent heat probably represents the work which has to be done in loosening the molecules of a solid when it becomes a liquid.

The latent heat of solids may be measured by means of the method of mixtures. Thus suppose W grams of a solid, of which the latent heat of fusion is L, at a temperature t_1 are placed in a calorimeter, the water equivalent of which and of its contents is w, and that the temperature of the calorimeter falls from t_2 to t_3. If s is the specific heat of the body in the solid state, s^1 the specific heat in the liquid state, and t_0 the melting-point of the body, then the heat absorbed by the body in being heated from t_1 to t_0 in the solid state, then melting at t_0, and finally rising from t_0 to t_3 in the liquid state, is

$$Ws(t_0 - t_1) + WL + Ws^1(t_3 - t_0),$$

while the heat lost by the calorimeter and its contents is

$$w(t_2 - t_3).$$

Equating these two quantities of heat, we get

$$L = \frac{w(t_2 - t_3) - Ws(t_0 - t_1) - Ws^1(t_3 - t_0)}{W}.$$

If the solid is originally at its melting-point, t_1 is equal to t_0, and no heat is used in raising the temperature of the solid, so that L is given by

$$L = \frac{w(t_2 - t_3) - Ws^1(t_3 - t_0)}{W}.$$

The following table gives the latent heat of fusion of some substances.

LATENT HEAT OF FUSION.

	$\frac{\text{Calories}}{\text{gram.}}$		$\frac{\text{Calories}}{\text{gram.}}$
Ice	80.02	Zinc	28.13
Sulphur	9.37	Lead	5.86
Paraffin	35.10	Silver	21.07
Benzene	30.85	Mercury	2.82

212. Bunsen's Ice Calorimeter.—Bunsen has utilised the change in volume, which takes place when ice is melted, to estimate the quantity of ice melted, and hence, knowing the latent heat of ice, to obtain a measure of the heat employed. His ice calorimeter consists of a glass test-tube A (Fig. 166), fused into a cylindrical glass bulb B. The lower part of this bulb is connected by a glass tube C, with a horizontal glass tube D, of fine bore, to which a scale is attached. The upper part of B is filled with distilled water, which has been well boiled to free it from dissolved air, the lower part and the side-tube being filled with mercury. By passing alcohol, which has been cooled in a freezing mixture, through A, a coating of ice is formed all round the lower part of A. The instrument is then packed round with melting snow, and left till the temperature of the whole apparatus comes to zero. To determine the specific heat of a substance, a known mass, heated to a temperature t, is dropped into A, and the amount of ice melted calculated from the distance the mercury recedes along the graduated tube. The instrument is often calibrated by introducing a known mass of water at a temperature t, and noting the number of divisions through which the mercury recedes, and then calculating the quantity of heat given to A, which causes the mercury to recede through one division.

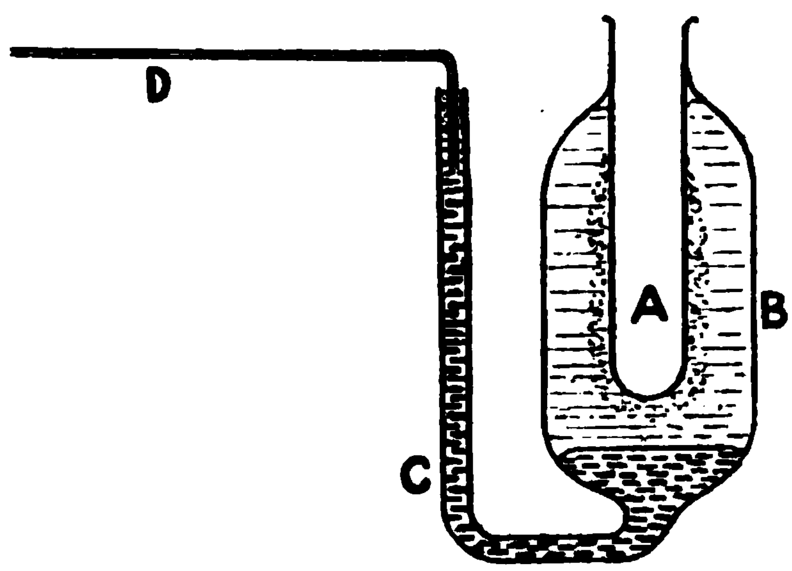


FIG. 166.

213. Boiling-Point.—When water is heated in a beaker the temperature gradually rises, and, unless the water has been very carefully freed from dissolved gas, as the temperature gets near 100° C. small bubbles are formed, mostly on the sides of the containing vessel. If the atmospheric pressure is 76 cm. of mercury, under ordinary circumstances, when the temperature reaches 100° C. bubbles are rapidly formed, and, rising to the surface, burst, and the temperature remains constant. The water is now said to boil. As has been mentioned in dealing with the upper thermometric fixed temperature, the temperature of boiling water depends on the pressure to which the water is subjected. Unless otherwise stated, it is usual to give the boiling-point of a liquid under the pressure of a standard atmosphere (§ 133).

The temperature of the *liquid* when ebullition takes place depends slightly on the nature of the containing vessel. The temperature of the vapour given off is, however, independent of the nature of the vessel, and hence, in determining the boiling-point of a liquid, the thermometer is usually placed in the vapour and not in the liquid itself.

The following table gives the boiling-point of some bodies under a pressure of one atmosphere :—

BOILING-POINTS.

Zinc	958° C.	Carbon dioxide	− 79° C.
Sulphur	444.5	Oxygen	− 183
Mercury	357	Air	− 192
Water	100	Nitrogen	− 194
Ethyl alcohol	78	Hydrogen	− 238
Ether	34.6		

214. Latent Heat of Vaporisation.—As in the case of the conversion of a solid into a liquid, so, in the case of conversion of a liquid into a vapour at the same temperature, heat has to be supplied. The quantity of heat that has to be supplied to one gram of the liquid, at the boiling-point, to convert it into vapour without changing its temperature or the quantity of heat given out by one gram of the vapour, at the boiling-point, when condensing to liquid at the same temperature, is called the *latent heat of vaporisation.* The boiling-point in the above definition is the temperature of ebullition under one standard atmosphere.

One of the simplest methods of determining the latent heat of vaporisation is that designed by Berthelot and shown in Fig. 167. The liquid to be experimented upon is contained in a glass flask D, down the centre of which runs a tube *ab*, open at both ends. The lower end of this tube is connected by a ground joint to a glass spiral S, which terminates in a small reservoir R, and an exit-tube open to the air. The spiral and reservoir are contained in a calorimeter which, to protect it against radiation, is itself contained in an outer vessel. The liquid in the flask is boiled by means of a ring burner B, the calorimeter being protected from

the heat by being covered with a slab of wood HH′. The liquid boils, and the vapour travels down the tube *ab*, is condensed in the spiral, and collects in the reservoir R. By this arrangement the condensation of the liquid, before it reaches the calorimeter, is avoided. The heat conducted to the calorimeter by the glass tube *ab* is allowed for by noticing the rate at which the temperature of the calorimeter rises before and after the experiment, when the flask is heated, but no distillation is taking place. The weight of liquid condensed is obtained by weighing the spiral and R before and after the experiment.

This apparatus is very convenient, for it only requires about 50 grams of the liquid, and the experiment only lasts three or four minutes. By its means Berthelot obtained as a mean the value 536.2 calories as the latent heat of vaporisation of water, while Regnault, using very elaborate apparatus, obtained 536.6 calories. Griffiths has recently obtained the value 536.63 calories (calories at 15°, see § 199).

Fig. 167.
(From Ganot's "Physics.")

The latent heat varies with the pressure, and therefore temperature, at which vaporisation takes place; and, according to Griffiths, the latent heat at a temperature t is, in the case of steam, given by the expression

$$L_t = 596.73 - 0.601t.$$

215. Joly's Steam Calorimeter.—Dr. Joly has invented a form of calorimeter in which the heat necessary to raise the temperature of the substance of which the specific heat is being measured, from a known temperature of about 20° to 100°, is obtained by determining the weight of steam which must be condensed to supply the necessary heat. The arrangement employed in measuring the specific heat of a gas at constant volume is shown in section in Fig. 168. The gas is contained in a copper sphere A, suspended by means of a fine wire C, from one arm of a delicate balance D, which is supported on a shelf above the apparatus. This wire passes through a small hole in the top of a copper vessel B, which is itself enclosed in a non-conducting box. Steam is admitted to the box B through the tube E, and that which is not condensed within the apparatus passes out through the tube F.

When the steam is admitted it condenses on the sphere A till the temperature reaches 100°, and the water formed by the condensation is collected in a thin catch-pan G, attached to the bottom of the sphere, and its weight is determined by putting weights on the opposite pan of the balance till equilibrium is again secured. A light metal shield H,

with a hole through which the suspending wire C passes, serves to protect the sphere from any drops of water produced by the steam condensing on the top of the vessel B. In order to prevent the condensation of steam on the wire C, where it passes through the hole in B, a spiral of fine platinum wire, I, is placed round the wire, but not touching it, and this spiral is heated by passing a current of electricity. In this way the portion of the wire passing through the hole is heated above 100°, so that no steam condenses on it.

Fig. 168.

In the best form of the steam calorimeter there is a sphere, &c., suspended from each of the arms of the balance, so that they are alongside each other in the vessel B. An experiment is then made, in which both the copper spheres are exhausted, and if they have exactly the same "water-value," the balance will remain in equilibrium after the admission of the steam. If the balance is deflected, weights are added till it comes back to equilibrium, and from the value of these added weights the difference in the water-value of the spheres can be calculated. One of the spheres is then filled with the gas to be experimented on under a pressure of about 40 atmospheres, and from the increase in the weight of the sphere the mass of the gas is obtained. The sphere is then placed in the calorimeter, the sphere attached to the other arm being still exhausted, and steam is admitted. The sphere containing the gas now condenses more steam than the empty one, since it requires some heat to raise the temperature of the enclosed gas. The weight w, which has to be added to the other side to produce equilibrium, is then equal to the weight of the water produced by the condensation of a weight w of steam, and the latent heat given out by this steam has heated the gas in the sphere from a temperature t, say, to 100°. Hence if M is the mass of the gas, and L the latent heat of steam, the specific heat (s) of the gas is given by

$$s = \frac{wL}{M(100-t)}.$$

The thermal value of the copper containing-sphere does not come in,

since this is compensated by the empty sphere attached to the other arm of the balance.

By means of this calorimeter, Joly has found the following values for the specific heat of some gases at constant volume, at a pressure of about 20 atmospheres :—

SPECIFIC HEAT AT CONSTANT VOLUME.

Air	0.1721
Carbon dioxide	0.1730
Hydrogen	2.402

216. Vapour Pressure.—If a small bubble of air is allowed to pass into the Torricellian vacuum of a barometer, the mercury column is depressed, and if a succession of bubbles are passed up, each will produce a depression. If, however, a small drop of a liquid, say ether, is introduced the column will be depressed, and the ether become entirely vaporised even at a temperature much below its ordinary boiling-point. If successive small drops of ether are introduced, it will be found that after a time the further addition of ether does not produce an additional depression, and that the ether no longer vaporises, but simply floats as a liquid on the top of the mercury column. If the space above the mercury be increased or decreased, by raising or lowering the barometer-tube in the cistern, it will be found that, so long as there is any liquid present, the *height* of the mercury column remains constant, but that the quantity of ether which vaporises varies with the space above the mercury. If the temperature is increased, more ether vaporises, and the mercury column becomes more depressed, and *vice versâ*.

The depression of the mercury column indicates that the liquid forms a vapour in the Torricellian vacuum, and that this vapour exerts a pressure on the upper end of the mercury column which partly balances the atmospheric pressure. The amount by which the column is depressed is a measure of this pressure which is called the *vapour pressure* of the liquid. When an excess of liquid is present, so that the vapour exerts its maximum pressure, and no more liquid will vaporise at the given temperature into the space under consideration, the vapour is said to be *saturated*. When, however, a given space contains some vapour, but if some more liquid were introduced some of it would vaporise at the given temperature, the vapour is said to be *unsaturated* or *superheated*.

The vapour pressure, or tension, as it is sometimes called, of a liquid depends on the temperature only. In the case of non-saturated vapours, Boyle's and Charles's laws apply approximately, the approxi-

mation being the better the further the vapour is removed from its saturation-point.

Suppose some liquid is contained within a cylinder which is closed by a freely moving piston, and that a pressure P acts on the outside of this piston. If the temperature of the liquid is below its boiling-point at the pressure P, the vapour tension will be less than P, and the pressure of the vapour on the inside of the piston will be less than that on the outside, so that the piston will rest on the surface of the liquid. As the temperature of the liquid is raised, the vapour pressure increases; when the vapour pressure is equal to the pressure P acting on the outside of the piston, this latter is in equilibrium. If the temperature rises ever so little more, the vapour pressure will be greater than P, and so the piston will be driven out, and vapour will be formed freely above the liquid. Now, exactly the same thing occurs when a liquid is heated in an open vessel, so that, when vapour is formed freely, the vapour pressure is equal to the pressure of the atmosphere acting on the surface of the liquid. Since when a liquid is vaporising freely it is said to boil, we have that when a liquid boils the vapour pressure at that temperature is equal to the external pressure to which the liquid is subject, while at temperatures below the boiling-point the vapour pressure is less than the external pressure. At temperatures above the boiling-point the *liquid* cannot exist, and the vapour is unsaturated.

217. Vapour Density.—In order to determine to what extent unsaturated vapours obey Boyle's and Charles's laws, the usual method is to determine the density, *i.e.* the volume of a gram of the vapour, at different pressures and temperatures. For temperatures lower than about 300°, the most convenient and accurate method for measuring the density of a vapour is one due originally to Gay-Lussac, but modified by Hofmann. A tube A (Fig. 169) about 80 cm. long, and having a bore of about 1 cm., is closed at one end and graduated, the volume from the closed end up to each division having been determined by weighing the mercury which filled it up to the division. This tube is surrounded by another tube, B, to which two side-tubes are attached. The tube A is filled with pure dry mercury, and then inverted in a dish E containing mercury. The vapour from some boiling liquid, contained in the

FIG. 169.

vessel F, is introduced at C, and escapes at D, and thus heats the inner tube to some uniform temperature. For instance, the following series of liquids boil under one atmosphere at the temperatures given, and hence would heat the apparatus to this temperature if they are used :—

	Deg. C.		Deg. C.
Carbon bisulphide	46.2	Bromobenzene	156.1
Ethyl alcohol	78.3	Aniline	184.4
Water	100.0	Methyl salicylate	222.9
Chlorobenzene	132.1	Bromonaphthaline	277

The liquid whose vapour is to be experimented upon is placed in a small glass bottle, which is shown full size at G, the difference in weight of the bottle empty and full giving the weight of the liquid, and hence also that of the vapour. The bottle is passed up the tube A, and when it reaches the Torricellian vacuum the stopper is driven out, and the contents, if too much liquid has not been taken, are completely vaporised. The volume of the vapour is read off on the tube, and the pressure to which it is subjected is obtained by subtracting the height of the column of mercury (corrected for change of density with temperature) from the barometric height, while the temperature is obtained by means of the thermometer T. At temperatures above about 100° it is necessary to take account of the vapour pressure of the mercury in the tube, which is added on to the pressure exerted by the vapour.

A convenient method of roughly measuring the vapour density of a body which at atmospheric pressure does not require a very high temperature to vaporise, is that devised by Victor Meyer. His apparatus is shown in Fig. 170, and consists of a glass bulb A connected to a straight stem B, about 70 cm. long, which is closed at the top by a cork E, and has a side delivery-tube C attached. This delivery-tube opens beneath the end of a graduated glass tube D, which is filled with water, and stands in a pneumatic trough. An outer tube F is partly filled with a liquid which can be heated to a temperature above the boiling-point of the substance to be tested. A weighed quantity of the substance contained, if it is a liquid, in a small stoppered glass bottle is introduced, and the cork E rapidly

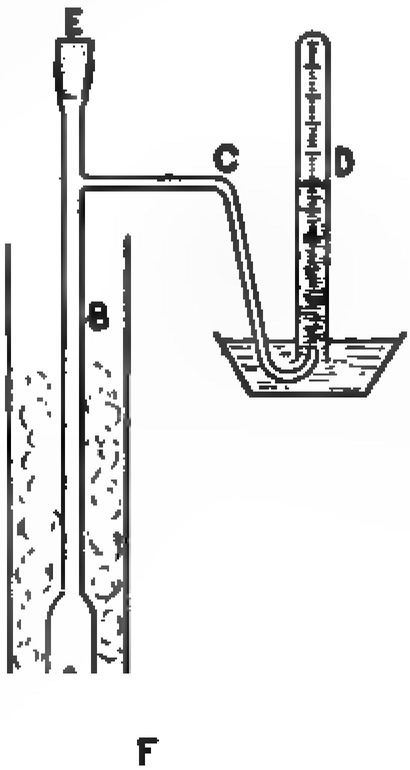


FIG. 170.

replaced. The substance vaporises within the bulb A, and in doing so drives some of the air out of B *through* C. The volume of this air will be equal to the volume of the vapour if it were at the temperature and pressure of the air in D. The reason is, that although the vapour displaces a volume of hot air from A equal to its own volume when at the temperature of the liquid in F, this air, being driven into the upper and cold part of the tube B, becomes cooled, and contracts according to Charles's law. Hence the volume of air actually expelled from the apparatus is equal to the volume which would be occupied by the hot air displaced by the vapour if it were cooled down to the temperature of the room. The volume of the air in D has to be reduced to standard pressure and temperature, corrections being applied for the vapour pressure of the water within D, and for the weight of the column of water in D above the surface of the water in the pneumatic trough. From this reduced volume of the vapour and the mass we can then calculate the density.

For high temperatures, a method due to Dumas is employed. A glass globe, the capacity of which is about half a litre, and having a neck drawn out to a fine point, is taken, and about 30 grams of the substance is introduced. The globe is then immersed in a bath of water, oil, or molten metal, at a temperature above the boiling-point of the substance, the end of the neck just projecting above the surface. The body (solid or liquid) is vaporised, and a jet of vapour spurts out of the neck of the globe, carrying with it the air contained in the globe. Directly the whole of the substance has vaporised, the jet of vapour escaping ceases, and the globe is now full of vapour at atmospheric pressure and at the temperature of the bath. The end of the neck is then sealed up by means of a blow-pipe. When the globe is cold it is weighed, it having been weighed previously before the introduction of the substance; the difference in weight gives the weight of the vapour, less the weight of an equal volume of air which has been driven out. The end of the neck is then broken off below the surface of some water, and since the vapour will have now condensed, the water is sucked up and fills the globe. Another weighing gives the weight of water contained in the globe, and hence the volume; and, knowing the density of air at the pressure and temperature of the first weighing, the density of the vapour can be calculated. For very high temperatures, Deville and Troost have replaced the glass globe by one of porcelain, using the vapours of sulphur, cadmium (815°), and zinc (930°), to heat the globe.

The density of a vapour having been measured at a temperature t, well above the boiling-point, and at a pressure p, the density it would have at the standard temperature t_0 and pressure p_0, supposing it could exist as a perfect gas under these conditions, is calculated by means of Boyle's and Charles's laws. Thus, if ρ is the observed

density, and ρ_0 the reduced density, we have, if we consider unit mass of the gas,

$$\frac{p}{\rho}=\frac{p_0}{\rho_0}\left\{1+a(t-t_0)\right\},$$

and hence

$$\rho_0=\frac{p_0\rho}{p}\left\{1+a(t-t_0)\right\},$$

where $a=0.00366$.

In the following table, the values of the density of some gases and vapours at 0° C. and under a pressure of a standard atmosphere are given, both in *c.g.s.* units (grams per cubic centimetre) and also in terms of the density of hydrogen taken as 2 :—

	Density.		Molecular Weight in terms of H=2.
	Grams per c.c.	In terms of H=2.	
Hydrogen (H_2)	0.00089551	2.0	2.0
Oxygen (O_2)	0.0142923	31.92	31.92
Nitrogen (N_2)	0.01257	28.1	28.01
Carbon dioxide (CO_2)	0.01977	44.6	43.89
Carbon monoxide (CO)	0.01251	27.9	27.93
Ammonia (NH_3)	0.00763	17.0	17.01
Chloroform ($CHCl_3$)	0.05431	121.3	119.08
Nitric oxide (NO)	0.01341	29.9	29.96

A consideration of this table will show that the density is proportional to the molecular weight, so that, when they are both measured in terms of hydrogen, the numbers as given in the third and fourth columns are identical. Since the molecular weight represents the weight (w) of a molecule, if N is the number of molecules in a cubic centimetre, and ρ the density in grams per c.c., we have

$$Nw=\rho, \text{ or } N=\rho/w.$$

But, as shown in the above table, ρ/w is constant for all gases ; hence N, or the number of molecules contained in unit volume of all gases under the same conditions of pressure and temperature, is the same. The above is Avogadro's law, and is of extreme utility in chemistry for determining the molecular weight of bodies which can be obtained in the form of a gas (*i.e.* vapours and gases).

The values of the molecular weights as deduced from the vapour density in the case of some bodies do not, at any rate at some temperatures, agree with the values deduced from the chemical behaviour of the body.

In the following table are given the values of the density (in terms

of H = 2) obtained at various temperatures for five of these anomalous bodies :—

	Temperature, Deg. C.	Density, H=2.	Molecular Weight, H=2.
Nitrogen tetroxide	29	85.8	$N_2O_4=91.9$
"	100	49.4	$NO_2=46.0$
"	135	46.2	
Phosphorus pentachloride	182	186.7	$PCl_5=207.7$
"	250	115.2	(Density of $PCl_3+Cl_2=104$)
"	300	105.4	
Iodine	448	252.4	$I_2=253$
"	940	220.9	(Density of $I+I=126.5$)
"	1470	146.1	
Acetic acid . . .	130	90.1	$C_2H_4O_2=59.86$
"	200	64.1	
"	300	60.1	
Sulphur	520	191.2	$S_2=63.96$
"	660	84.6	
"	1040	64.7	
"	1400 (about)	63.5	

In the case of nitrogen tetroxide, the vapour density, at a temperature of 135°, corresponds to the molecule NO_2. At lower temperatures the density corresponds more nearly to the molecule N_2O_4. Thus it would seem that, as the temperature is raised, each molecule of N_2O_4 splits up into two molecules of NO_2. The values obtained for the vapour density of phosphorus pentachloride in the same way show that, even at 182°, some of the molecules of PCl_5 have split up into a molecule of PCl_3 and a molecule of Cl_2, while at a temperature of 300° this dissociation is almost complete. At a temperature of 448° the iodine molecule consists of two atoms, while at a temperature of 1470° each of these molecules has split up into two molecules. In the case of acetic acid and sulphur at low temperatures, we have the opposite phenomenon to dissociation taking place, namely, the association of the molecules to form complete molecules. Thus at a temperature of 520° the molecule of sulphur appears to be S_6. At high temperatures these associated molecules break down, and we get the normal vapour density.

218. The Measurement of Vapour Pressure.—The determination of the maximum vapour pressure which a liquid possesses at a given temperature, or the vapour pressure of the saturated vapour, can be performed for ordinary temperatures by means of Hofmann's apparatus for vapour density (Fig. 169). In this case liquid is introduced into the tube till it ceases to evaporate, and a thin layer floats on the top of the

mercury column, and the pressure is obtained by measuring the height of the mercury column and the barometric height. The height of the mercury column has to be reduced to what it would be if the temperature of the mercury were 0°, and at temperatures above 100° a correction has to be applied for the vapour pressure of mercury. The chief objection to this method is that the layer of liquid on the top of the mercury column alters the capillary constant of the mercury and glass surface, and necessitates a correction of doubtful amount.

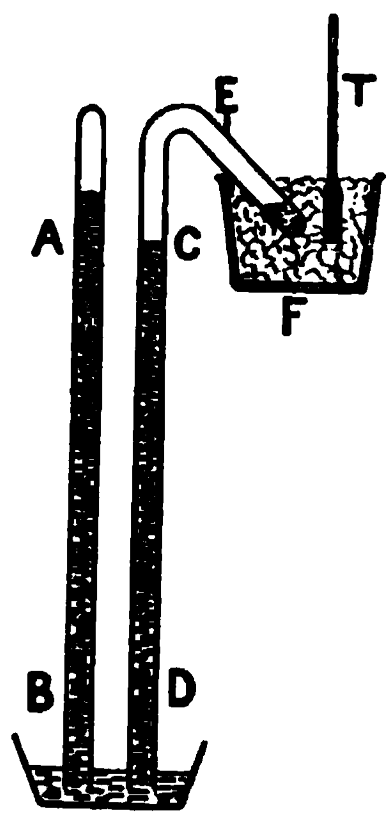


FIG. 171.

For low temperatures Gay-Lussac used the arrangement shown in Fig. 171. The liquid was introduced into the tube DCE, which is an ordinary barometer tube with the end E bent round, so that it can be immersed in a freezing mixture. The other tube, AB, acts as a barometer for measuring the atmospheric pressure. Under these circumstances the vapour pressure in the upper part of DCE corresponds to the maximum vapour pressure at the temperature of the coldest part, *i.e.* E. That this must be is evident, if we consider two bulbs, A and B (Fig. 172), connected by a tube C, the one bulb, A, being at a temperature t_1, and the other at a lower temperature t_0. Suppose we start with some liquid in both bulbs, then the vapour pressure of the liquid in A will be p_1, say, and that in B p_0, where $p_1 > p_0$. If then we consider a piston placed in C, the pressure on the left would be greater than that on the right, and the piston would be driven over towards B, thus increasing the pressure of the vapour in B, and therefore causing some of the vapour to condense, for B is already full of vapour saturated at t_0. The pressure in A would at the same time be reduced, and hence the vapour would no longer be saturated, and some of the liquid would be vaporised. This action goes on although the piston we have imagined does not exist, and will continue till all the liquid in A, the bulb at the higher temperature, has been evaporated, when the vapour pressure throughout becomes p_0, *i.e.* corresponds to the temperature of the coldest part of the enclosure. The vapour in A is then no longer saturated, while in B it is saturated. In Gay-Lussac's apparatus, therefore, the vapour pressure, as measured by the difference in height of the mercury in columns AB and CD, corresponds to the temperature of the bath F. There is a further advantage, in that there is no liquid to affect the capillarity of the mercury in the tube CD.

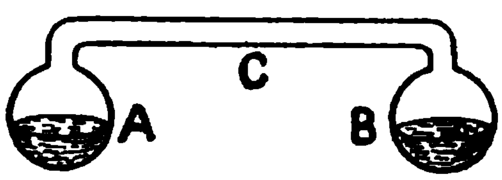


FIG. 172.

For pressures greater than atmospheric pressure, Regnault designed a form of apparatus in which the vapour itself kept the temperature

constant while an experiment was being made, and which could also be used for pressures less than an atmosphere. The liquid is enclosed in an air-tight metal vessel A (Fig. 173), from which an inclined tube leads to a copper globe B. This globe is surrounded by a water bath to keep its temperature constant, and can be connected by means of a three-way cock, F, to a compressing or exhausting pump, and to a mercury manometer. The liquid in A is boiled, and the vapour passes up into the inclined tube, where it is condensed by a stream of cold water which passes through the condenser C, and then flows back into the boiler A. The temperature of the vapour over the boiling liquid is given by four thermometers T, which are placed in four iron tubulars, which are closed at the bottom, and contain mercury. Since a liquid boils when its temperature is such that the maximum vapour pressure is equal to the pressure to which it is subjected, the manometer gives the vapour pressure corresponding to the temperature as given by the thermometers T. The pressure in the globe having been adjusted to the required value, either greater or less than the atmospheric pressure, and the flask heated, boiling soon starts, and in a very short time the temperature becomes absolutely constant, and remains so as long as is required. The manometer and the thermometers T having been read, the pressure is altered by means of the pump, and when ebullition has continued for a few minutes, the readings for the new pressure are taken, and so on. In this arrangement, when the steady state has been reached, the heat supplied by the burner is employed in supplying the latent heat of vaporisation of the liquid. The vapour then passes to the condenser, where it parts with its latent heat and again becomes liquid, and returns to the vessel A, running down the inclined tube. The pressure does not rise, since as much vapour is condensed during each second as is produced. If the supply of heat is increased, the rate at which the vapour

is produced is also increased, and the only effect of this is that the vapour is able to pass a little further up the condenser before it is all condensed; but since the condenser is always made so long that the vapour never reaches the further end, no vapour passes over to the globe B. Thus the rate at which the vapour is condensed is increased, and remains equal to the rate at which it is vaporised, so that the pressure does not alter.

The following table gives the maximum vapour pressure of four liquids at different temperatures :—

VAPOUR PRESSURE IN CM. OF MERCURY.

Temperature.	Carbon Bisulphide.	Ethyl Alcohol.	Water.	Mercury.
Deg. C.	Cm.	Cm.	Cm.	Cm.
0	12.8	1.22	.46	0.00002
5	...	...	.65	...
10	19.8	2.38	.91	0.00005
15	...	...	1.27	...
20	29.8	4.40	1.74	0.0001
30	43.5	7.81	3.15	0.0003
40	61.8	13.37	5.49	0.0008
50	85.7	22.0	9.20	0.0015
60	...	35.0	14.89	0.0029
70	...	54.1	23.33	0.0052
80	...	81.2	35.49	0.0092
90	...	118.7	52.55	0.0160
100	...	169.2	76.00	0.0270
150	...	736.9	358.1	0.2684
200	...	2218.2	1168.9	1.7015

In Fig. 174, the curve showing the connection between the vapour pressure of water and the temperature has been plotted. This curve divides the diagram into two regions, in one of which the conditions are such that the water can only exist as an unsaturated vapour, and in the other only as a liquid, while along the curve we may have the liquid and vapour existing simultaneously, *i.e.* the vapour is saturated. For suppose we had some water enclosed in the Torricellian vacuum of a barometer tube, the pressure being 20 cm. of mercury and the temperature 80°. The conditions are represented by the point A on the diagram. From the curve we see that the maximum vapour pressure corresponding to a temperature of 80° is 35.5 cm., so that the vapour is unsaturated. If now the pressure on the vapour is increased, the temperature remaining constant, the conditions the vapour passes through are represented by the vertical line AB. When the point B on the curve is reached, the pressure is equal to the maximum vapour pressure, and if the pressure is increased beyond this point the vapour will condense into a liquid. In

the same way if, starting from A, we keep the pressure constant, reducing the temperature, the changes are indicated by the straight line AC. When the point C is reached, *i.e.* the temperature falls to 66°, the vapour

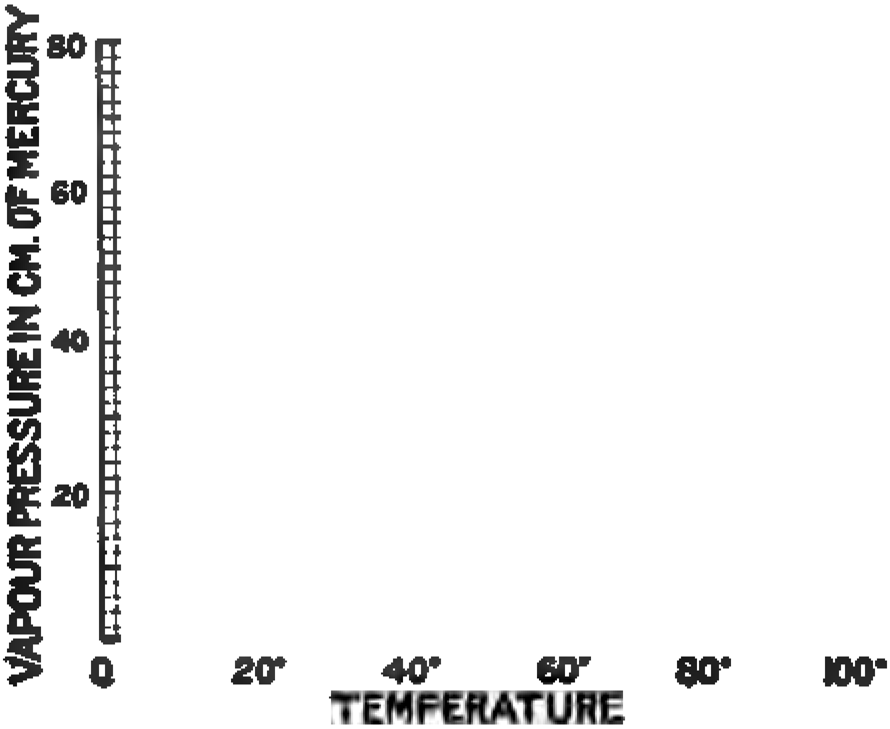


Fig. 174.

will be saturated. Any further fall of temperature will be accompanied by the condensation of the vapour into a liquid. Hence, corresponding to all points on the diagram to the right and below the curve we have vapour, and to those on the left and above we have liquid.

219. Mixtures of Vapours and Gases.—In the previous sections we have considered the formation of vapour in a space which was free from gas, we have now to consider the formation of vapour when the space over the liquid already contains a gas such as air. Dalton, who first investigated this question, found that if some liquid is introduced into an enclosure which contained a gas at a pressure H, then the pressure in the enclosure rises, and if the whole of the liquid does not evaporate, *i.e.* if there is enough liquid to saturate the space, the final pressure $H+h$ is such that h represents the maximum vapour pressure of the liquid at the temperature of the experiment. Hence, as far as the vapour is concerned, a space filled with a gas behaves as a vacuum, the only difference being that in a vacuum the space becomes saturated almost immediately after the introduction of the liquid, while when a gas is present the evaporation of the liquid is much slower, and hence it takes some time to saturate the space. From the results of his experiments Dalton enunciated the two following laws, which are known by his name: (1) The pressure exerted by, and the quantity of, a vapour which

saturates a given space are the same for the same temperature, whether this space is filled by a gas or is a vacuum. (2) The pressure exerted by a mixture of a gas and a vapour, of two vapours or of two gases, is equal to the sum of the pressures which each would exert if it occupied the same space alone.

In order to verify the accuracy of Dalton's laws, Gay-Lussac used the apparatus shown in Fig. 175. The glass tube A is closed above by a special form of tap C, in which the barrel, instead of being pierced completely, has only a small recess made at one point. The lower end of A is connected to a side-tube B, which acts as a manometer, and has a tap D by which mercury can be withdrawn. When the stopcock C is turned with the recess upwards, this becomes filled with any liquid placed in E, and when the stopcock is turned round through 180° the liquid filling the recess is discharged into A. The position of the mercury in the two limbs is noted before the introduction of the liquid, then enough liquid is introduced to saturate the space, and mercury poured into B till the level of the mercury surface in A comes back to its original position, so that the gas now occupies the same volume it did before the introduction of the liquid. The difference in the levels of the mercury in B before and after the introduction of the liquid gives the pressure exerted by the vapour, and this pressure will be found to be equal to that exerted by some of the same liquid when introduced into a Torricellian vacuum at the same temperature.

FIG. 175.

It would seem *à priori* that Dalton's law can only be an approximation, for otherwise it would mean that, by introducing a sufficiently large *number* of different *kinds* of liquids into the same space, we could produce as great a pressure as we please, a result that is unlikely to be true. Regnault, who investigated the pressures of mixtures of gases and vapours, found that in the case of vapours formed in air and in nitrogen, the two gases he tested, the vapour pressure was very slightly less than in a vacuum. The differences, however, were so small that he considered they might be due to the condensation which always takes place on the glass walls of the apparatus, so that Dalton's laws may be true in the cases of mixtures of gases and vapours. In the case of mixtures of two vapours, Magnus, and subsequently Regnault, found that the pressure of the vapour of a mixture of two or more liquids which do not mutually dissolve one another is equal to the sum of the pressures they would each exert separately, but that when the liquids mix the vapour pressure of the mixture is less than the sum of the vapour pressures of the constituents.

Experiments by Andrews show that, in the case of a mixture of two

gases, Dalton's law only holds if the gases are far removed from their point of liquefaction, *i.e.* are practically in the condition of perfect gases.

220. Humidity of the Atmosphere—Hygrometric State.—The atmosphere consists of a mixture of oxygen and nitrogen in a practically constant ratio, together with some small quantities of other gases, and with a very variable amount of aqueous vapour. The maximum quantity of aqueous vapour which a given volume of air can contain is, of course, equal to the mass of vapour this volume would contain when filled with saturated water vapour at the given temperature. Ordinarily, however, the air contains less aqueous vapour than would saturate it, and the ratio of the pressure (f) exerted by the aqueous vapour actually present to the maximum vapour pressure (F) at the actual temperature is called the *humidity*, *relative humidity*, or *fraction of saturation* of the air. Our sensations as to the dryness or dampness of the air depend on the above ratio, and not on the actual quantity of vapour present in the air. Thus on a cold winter's day, when the air is saturated at a temperature of say 5°, the air feels very damp, while if the temperature had been 15° the same quantity of moisture would not nearly saturate the air (the humidity would be $.65/1.27 = 0.51$; see table of vapour tension of water, p. 259), and it would feel comparatively dry.

The humidity (f/F) may also be expressed as the ratio of the weight (w) of vapour actually present[1] in a given volume of air to the weight (W) which would saturate the same volume of air at the given temperature. Since unsaturated vapours obey Boyle's law, the weight of the vapour in a given volume is proportional to the pressure exerted by the vapour. Hence

$$\frac{w}{W} = \frac{f}{F}.$$

If air containing aqueous vapour is cooled, a temperature will eventually be reached such that the vapour saturates the space, and any further cooling will cause condensation of some of the vapour into water. This temperature is that at which the air would be saturated (*i.e.* have a humidity 1) if it contained the same quantity of water that it has at the original temperature, and is called the *dew-point.*

If t_1 is the actual temperature of the air and t_0 the dew-point, then, from a table giving the quantity of water vapour in unit volume of saturated air at the different temperatures, we can obtain W, the weight of water in unit volume saturated at t_1, and w that in unit volume saturated at t_0. But w is the weight of water actually present in the air, since we have supposed it cooled down to the dew-point without loss or gain of moisture. Hence the hygrometric state w/W can be obtained from a knowledge of the dew-point, and of the actual temperature of the air.

[1] The mass of aqueous vapour present in a cubic metre of air is often called the absolute humidity of the air.

The diagram given in Fig. 174 will assist in making this clear. Suppose (although, of course, such a high temperature would not occur in the open air) that the temperature of the air and the vapour pressure of the water present are represented by the point A, so that at a temperature of 80° the vapour pressure is equal to 20 cm. of mercury. The maximum vapour pressure corresponding to a temperature of 80° is 35.5 cm. of mercury. Hence the hygrometric state corresponding to the point A is 20/35.5. Now, if the air is cooled down, we shall travel along the line AC, but when the point C is reached, that is, at a temperature of 66°, the air will be saturated, and the deposition of dew will commence. The temperature corresponding to the point C will therefore be the dew-point. If, then, by experiment we determine the temperature of the dew-point, we can, from such a curve, or from a table of the vapour pressure of water, determine what is the maximum vapour pressure at the dew-point, and this is the actual vapour pressure present. Also, by observing the temperature of the air, we can in the same way obtain what would be the maximum vapour pressure at this temperature, and the ratio of these two numbers is the hygrometric state.

221. Hygrometry.—Hygrometers are instruments for measuring the hygrometric state of the air, and may be divided into three classes: (1) Those in which the dew-point is determined, called dew-point hygrometers; (2) those in which the actual weight of moisture contained in a measured volume of air is determined, called chemical hygrometers; and (3) wet and dry bulb hygrometers.

The most commonly used form of dew-point hygrometer is that devised by Regnault. This instrument consists of two glass tubes E and D (Fig. 176), the lower ends of which are closed by thin silver thimbles. They are each closed at the top by a cork, which supports a delicate thermometer (T and *t*).

FIG. 176.
(From Ganot's "Physics.")

Through the cork in D a glass tube A also passes, the end reaching nearly to the bottom of the thimble. The tube AD is connected by means of the tubulure, which fixes it to the stand, and an india-rubber tube with an aspirator G. Some ether is placed in the thimbles, and after the instrument has had time to reach the temperature of the air, the two thermometers are read, giving the temperature of the air, t_1. The aspirator is now started, and draws air through the tube A into the instrument. This air bubbling through the ether causes evaporation, which cools the ether and thimble, which in turn cools the air in its immediate vicinity. When a film of dew is deposited on the thimble D, indicating that the dew-point has been reached, the aspirator is stopped, and the temperature of the thermometer T read. It is again read when the dew disappears from the thimble, and the mean of these two readings gives the dew-point t_0.

In the chemical hygrometer a known volume of air is drawn, by means of an aspirator, through a series of tubes containing substances, such as anhydrous calcium chloride or phosphorus pentoxide, which readily absorb moisture. From the difference in the weight of these tubes before and after the passage of the air and the volume which has passed, the absolute hygrometric state of the air (w) can be obtained, and W can be got from tables, if the temperature of the air is taken.

The wet and dry bulb hygrometer depends for its action on the fact that the drier the air is, the more rapid will be the evaporation from a wet body exposed to the air. Since evaporation requires the supply of heat (latent heat of evaporation), it follows that the extent to which a wet body is cooled by evaporation will depend on the hygrometric state of the surrounding air. Two similar thermometers are fixed on a stand, the bulb of one of them being covered with muslin kept moist by means of a piece of lamp-wick which dips in a vessel of water. Unless the air is saturated, evaporation will take place from the muslin, and hence the wet bulb thermometer will indicate a lower temperature than the other, the difference being greater the greater the evaporation, that is, the drier the air. By comparing the readings of the wet and dry bulb thermometers with the humidity, as obtained by other hygrometers, a table has been drawn up, by means of which, from the reading of the dry bulb thermometer, and the difference between the dry and wet bulb thermometers, the dew-point can be obtained. The indications of this instrument are, however, considerably influenced by its environment, also by the action of draughts, &c.

222*. Effect of the Curvature of the Surface on the Vapour Pressure.—The form of the surface separating a liquid from its saturated vapour has an influence on the vapour pressure, which Lord Kelvin was the first to point out, and which has important applications in explaining the condensation of vapour into liquid in such cases as occur in clouds.

Suppose we have some liquid, such as water, contained within a vessel C (Fig. 177), from which all air has been exhausted, so that we have only to do with the liquid below and its vapour above. Further, let a fine capillary tube AD of radius r dip in the liquid. If the liquid wets the glass it will rise in the capillary, and let the height of the curved surface A above the plane surface B be h.

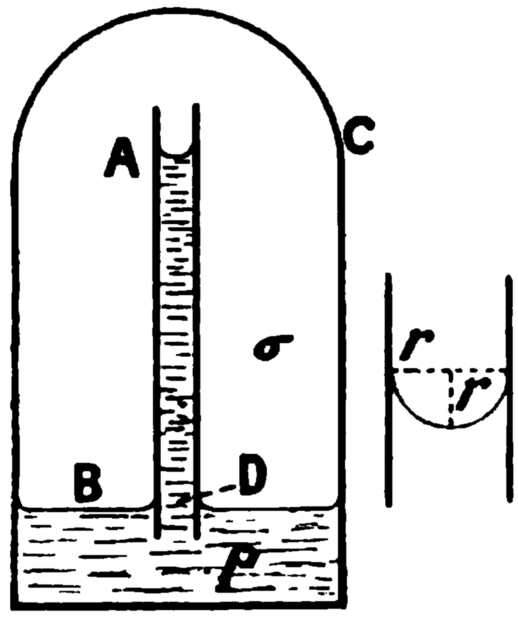


FIG. 177.

Now the pressure within the vapour at the level of A will exceed the pressure at the level of B by the weight of a column of *vapour* of height h, or, if σ is the density of the vapour, by σhg dynes per square centimetre. Hence, if the whole is at the same temperature, and if the vapour pressure at the concave surface A is the same as at the plane surface B, when the pressure at B is equal to the vapour pressure at the existing temperature, the pressure at A will be less than the vapour pressure, and so evaporation will still take place from the surface A. This would involve a continuous circulation of the liquid up the tube, for the height h depends on the surface tension, and must remain constant. Such a continuous circulation could, theoretically, be made to do external work, say by turning a small turbine placed in the tube; and since the temperature would remain constant, we should thus manufacture energy, which is contrary to the law of the conservation of energy. We therefore conclude that the liquid and its vapour must be in equilibrium both at A and at B, or that the vapour pressure, p, at the plane surface must be greater than that, C, at the concave surface A by an amount equal to the weight of a column of the vapour of height h, or

$$p - \mathrm{C} = \sigma hg.$$

If the density of the liquid is ρ, the weight of the column of liquid of height h is ρgh. The difference in pressure between the surface A and a point D within the tube on a level with the surface is equal to the weight of the column of the liquid of density ρ, less the difference of pressure between A and B, due to the weight of an equal column of the vapour of density σ. Thus the difference of pressure between A and D is

$$gh(\rho - \sigma).$$

If the liquid wets the tube, so that the angle of contact is 180°, it has been shown in § 160 that the difference in pressure between A and D is equal to

$$\frac{2T}{r},$$

where T is the surface tension of the liquid-vapour surface.

Hence, equating the two values we have obtained for the difference of pressure, we get

$$\frac{2T}{r} = gh(\rho - \sigma),$$

or, substituting for h its value $(p-c)/\sigma g$,

$$\frac{2T}{r} = \frac{(p-c)(\rho-\sigma)}{\sigma},$$

or

$$p - c = \frac{2\sigma T}{r(\rho - \sigma)}.$$

Now the curved surface of the liquid is, as shown at the side, very nearly a hemisphere of radius r, and we see from the above expression that the decrease of vapour pressure with curvature is inversely proportional to the radius of the spherical surface. If, instead of being concave, the surface had been convex, such as is the case in a raindrop, the vapour pressure at the curved surface would be greater than that at a plane surface, and this increase would *increase* with the decrease in the radius, r, of the drop. Thus, in the case of very small drops, the vapour pressure may be very considerably greater than that corresponding to a plane surface at the same temperature. The result is that although the air may be saturated, as measured in the ordinary way with a plane surface, very small drops, so far from increasing in size by the condensation of vapour, are actually evaporating.

The above reasoning explains why it is that if air is perfectly free from suspended solid matter, or dust, it may be cooled to a temperature considerably below the dew-point, without the formation of drops of water or mist. A very small drop—and at first, in such a dust-free air, all the drops must be small—will have a high vapour pressure, and will again evaporate. If, however, there is dust in the air, the dust particles will act as nuclei, so that the water which condenses first on them, instead of being in the form of an excessively small spherical drop, may be spread out into a surface of comparatively small curvature, so that re-evaporation will not take place. The formation of large drops is also explained, for the vapour tension at the surface of a small drop will be greater than that at the surface of a larger drop, and hence evaporation will take place from the small drops, and condensation on the large.

223. Sublimation.—Hitherto we have exclusively considered the passage of a solid to the liquid state, and that of a liquid to the gaseous state. Under certain conditions it is possible, however, for a solid to pass directly into the gaseous state without passing through an intermediate liquid condition. This change from solid to vapour is called *sublimation*, and is very clearly marked in the case of camphor and iodine. These bodies, when gently heated, readily pass into vapour, although the temperature has not been sufficiently high to melt them.

Although to a much less marked degree, ice exhibits the same phenomena. Thus at a temperature of $-1°$ C. the vapour tension of ice amounts to 0.42 cm. of mercury, and a piece of ice kept at this temperature will sublime till the pressure of the vapour in the surrounding space is 0.42 cm. of mercury, when equilibrium will be set up.

224. The Triple Point.—In Fig. 174 we have given the curve of maximum vapour pressure for a liquid (water), or, in other words, the boiling-point for different pressures. This curve gives the pressure corresponding to any temperature at which both the liquid and the vapour can exist in contact one with the other without their relative proportions altering—*i.e.* they are in stable equilibrium—and is called the *steam line.*

As has been seen in § 210, the melting-point of a solid depends on the pressure, so that a similar curve to the steam line can be drawn, giving the melting-point at different pressures. Such a curve will indicate the pressure corresponding to any temperature to which a mixture of ice and water must be subjected, in order that the two states may be in stable equilibrium. This curve is called the *ice line.* Finally, we may have, as has been mentioned in the previous section, a solid in stable equilibrium with its vapour, and may therefore draw a third curve showing the pressures at which, under various temperatures, a solid and its vapour can exist simultaneously. This curve is called the *hoar-frost line.*

The general form of these curves for water is shown in Fig. 178. The three curves meet at the point P, which is called the *triple point.* Since the steam line gives the conditions under which the vapour and liquid may exist simultaneously, the ice line those under which the liquid and solid may exist simultaneously, and the hoar-frost line those under which the vapour and solid may exist simultaneously, it is obvious that at the triple point all three, solid, liquid, and vapour, can coexist in stable equilibrium. The ice line in the case of water, which expands on solidifying, so that increase of pressure lowers the melting-point, slopes downwards towards the right. Since, however, the lowering per atmosphere increase of pressure is only 0°.0075, the slope is too small to be indicated on the figure. In the case of a body like paraffin, which contracts on solidifying, the ice line would slope downwards and towards the

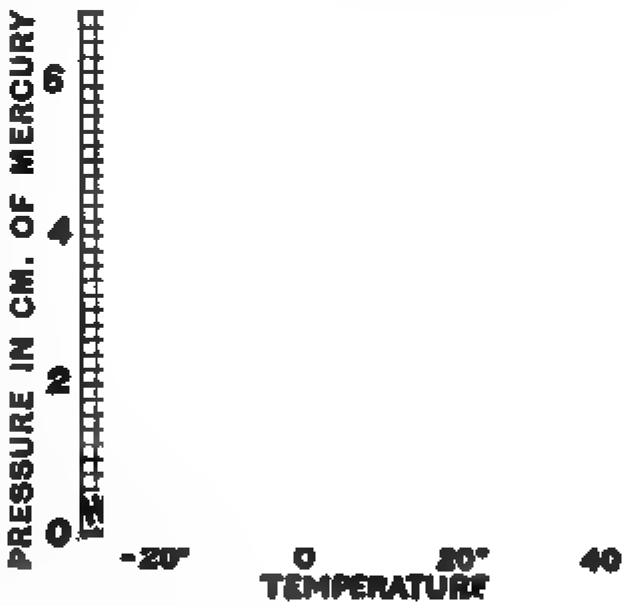


Fig. 178.

left. The triple point for water corresponds to a pressure of 0.046 cm. of mercury, and a temperature a very little above 0°.

225. Freezing-Point of Solutions—Cryohydrates.—It has long been known that the freezing-point of sea water is lower than that of pure water, and generally that the presence of a salt dissolved in water lowers the freezing-point. Of late years, however, great attention has been directed towards the effect of a dissolved salt on the freezing-point of the solvent, and the results are of very great interest, both from a physical and a chemical standpoint.

The first to make anything like a complete investigation of this subject was Raoult, and he found that the depressions produced by equi-molecular quantities of different substances dissolved in the same solvent were approximately the same, so long as the solutions were not too concentrated. By equi-molecular quantities is meant quantities of the different substances proportional to their molecular weights, so that the solutions contained equal numbers of molecules of the dissolved substances in the same volume. For fairly dilute solutions the depression is proportional to the quantity of salt dissolved. In the following table, the molecular depressions are given, *i.e.* those which would be produced if the molecular weight in grams of a body was dissolved in 100 grams of the solvent. These values are calculated, on the supposition that the depression is proportional to the concentration, from experiments made on much more dilute solutions, although with such concentrated solutions this proportionality no longer exists, and, even if it did, it would in many cases be impossible to obtain such concentrated solutions at such low temperatures. It is, however, convenient to reduce all results to some standard number of molecules of the dissolved substance to a given volume of the solvent, and the molecular weight in grams is in many ways a convenient number. The same kind of convention is employed when stating the density of a vapour, in that the density is given for a temperature of 0° and a pressure of a standard atmosphere (§ 217), although in most cases, under these conditions, the vapour would have condensed to a liquid.

MOLECULAR DEPRESSIONS FOR SOLUTIONS IN ACETIC ACID.

Chloroform	38°.6
Carbon bisulphide	38.4
Ether	39.4
Formic acid	36.5
Sulphur dioxide	38.5
Glycerine	36.2
Ethyl alcohol	36.4
Sulphuric acid	18.6
Hydrochloric acid	17.2
Magnesium acetate	18.2

MOLECULAR DEPRESSIONS FOR SOLUTIONS IN BENZENE.

Chloroform	50°.4
Carbon bisulphide	49.7
Ether	49.7
Ethyl alcohol	25.3
Formic acid	23.2
Acetic acid	25.3

MOLECULAR DEPRESSIONS FOR SOLUTIONS IN WATER.

Ethyl alcohol	17°.3
Cane sugar	18.5
Acetic acid	19.0
Hydrochloric acid	39.1
Sulphuric acid	38.2
Sodium chloride	35.1
Calcium chloride	49.9

It will be seen from the above numbers that, for any given solvent, the values of the molecular depression approximate to one of two constant values, one of these values being half the other. If we suppose that the depression is proportional to the number of dissolved molecules, and independent of the nature of the molecules (Van't Hoff's theory), the lower value of the molecular depression may be due to the fact that, in some cases, the molecules have formed into aggregates of two ordinary molecules, so that in the solution the molecular weight is doubled; or the higher value may be due to the splitting up or dissociation of the molecules when in solution. We shall refer to this question later, when we consider electrolytic conduction (Book V. Part VIII.). Assuming that the molecular depression is a constant, if we know its value for any solvent we can deduce the molecular weight of a body by observing the depression in the freezing-point it produces when dissolved in that solvent.

The most usual form of apparatus for determining the freezing-point is that designed by Beckmann, and shown in Fig. 179. The solution to be examined is placed in a glass test-tube A, which is surrounded by another tube B, with an air space between, the whole being placed in a glass beaker C. A freezing mixture is placed in C, and the temperature of the solution, as indicated by the thermometer T, is watched. It is generally found that the solution can be cooled down slightly below its freezing-point without ice forming. On stirring with the platinum wire E, small crystals of ice are formed,

FIG. 179.

and the temperature rises to a certain point, and then becomes stationary. This temperature is the freezing-point of the solution. The rise in the temperature is brought about by the latent heat of fusion of the small quantity of ice formed.

When a dilute solution in water is frozen, at first pure ice solidifies out, and on this account the concentration of the remaining solution increases, and the freezing-point becomes lower and lower. If the process is continued, a stage will at length be reached when the remaining solution is saturated at the existing temperature. Any further cooling will separate more ice, and hence, as the solution is already saturated, some of the salt must be deposited in the solid state; and since this deposit of the salt keeps the concentration of the solution constant, the temperature will not change till the whole of the water and the dissolved substance are solidified, one as ice, the other as the salt. That they are deposited separately, and not in chemical combination, seems to be indicated by the fact that the ice may be dissolved out by alcohol, leaving a skeleton of solid salt. These combinations of the solvent and dissolved substance, both in the solid state, have been regarded as definite chemical compounds, and as such were called cryohydrates.

In Fig. 180, the ordinates represent the percentage of common salt (sodium chloride) present in a solution in water, and the abscissæ represent temperatures. The curve AB shows the freezing-point of solutions of different strengths, that is, the temperature at which solids begin to separate from the solution. In the portion BP of this curve, the solid which separates first is pure ice, but at the point P salt also begins to separate out. The curve CD represents the quantity of common salt which will form a saturated solution at different temperatures. This solubility curve cuts the freezing-point curve at the point P, which corresponds to a temperature of $-22°$ C., so that for all the points in the curve AB, from B to P, the solution is unsaturated. At P, however, the solution is saturated, and hence at P salt commences to separate from the solution in the solid state. The strength of the solution, when both salt and ice are separated, is 23.8 parts by weight of salt to 76.2 parts of ice, and this is the so-called cryohydrate of sodium chloride. It is interesting to note that it is only possible to have liquid solutions of sodium chloride in states corresponding to the portion of the diagram (Fig. 180) included between the lines DP, PB. We cannot, of course, have stable solutions corre-

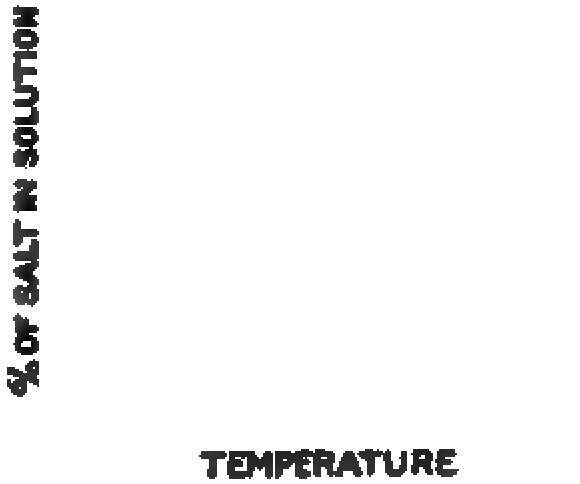


Fig. 180.

sponding to points above the solubility curve CD. Supersaturated solutions are unstable, for this would involve a solution containing a larger amount of dissolved salt than a saturated solution. To show that we cannot have a solution in the state represented by a point to the left of the freezing-point curve AB, suppose that we start with a solution in the state represented by the point E, that is, with a 5 per cent. solution at a temperature of 0°. If this solution is cooled, the change will be shown by the horizontal line EG. When the solution is cooled down to a temperature of $-3°$, that is, when we arrive at the point F, where the line EG cuts the freezing-point curve, *ice* will be separated, and hence the concentration of the solution will increase. Further cooling will thus cause us to traverse the curve FP, so that we shall not be able to get the solution into the condition corresponding to any point to the left of the curve AB. After the point P has been reached, the ice and salt will be deposited together in constant proportion, and the temperature will remain constant till the whole has solidified. When solidification is complete, the temperature, if the cooling is continued, will again fall, and, since the percentage of salt is now invariable, we shall travel along the horizontal straight line PQ, but we shall no longer be dealing with a solution, but with a mixture of ice and salt.

226. Heat of Solution — Freezing Mixtures.—When a body, whether it is a solid, a liquid, or a gas, is dissolved in a liquid, both being at the same temperature, there is in general a change in temperature indicating either an absorption or evolution of heat. The quantity of heat absorbed or liberated by the solution of one gram of a substance, in a quantity of the solvent so large that further dilution does not produce any further appreciable thermal change, is called the heat of solution of the substance.

In the following table, the heats of solution of some substances in water at a temperature of about 18° C. are given :—

HEAT OF SOLUTION.

Substance.	State.	Heat of Solution in Calories.
Chlorine	Gas	+68.9
Carbon dioxide	"	+134.0
Ammonia	"	+495.6
Hydrochloric acid	"	+476.1
Ethyl alcohol	Liquid	+55.3
Acetic acid	"	+7.02
Sulphuric acid	"	+182.5
Potassium hydroxide (KHO)	Solid	+223.3
Sodium chloride	"	−18.22
Potassium chloride	"	−59.7
Mercuric chloride	"	−12.2
Silver chloride	"	−110

In the above table, a plus sign indicates that solution is accompanied by evolution of heat. It will be noticed that while gases and liquids always give an evolution of heat on solution, solids sometimes give an evolution and sometimes an absorption.

By using comparatively large quantities of solid substances, for which the heat of solution is negative, dissolved in a limited quantity of water, very considerable falls of temperature can be produced. This is illustrated in the following table :—

FREEZING MIXTURES.

Substance.	Parts dissolved in 100 parts of Water.	Temperature.	
		Before Solution.	After Solution.
Sodium nitrate	75	+13.2	−5.3
Calcium chloride (cryst.)	250	+10.8	−12.4
Ammonium nitrate	60	+13.6	−13.6

Such solutions of salts in water are sometimes used to produce cold, and are called freezing mixtures. More efficient freezing mixtures may be employed, in which the lowering of temperature is produced on account of the change of state of one or both of the constituents. The thermal changes which go on in many of these mixtures are, however, very complicated. We may take, as a somewhat simple example, the case of a mixture of ice and sodium chloride. Suppose that powdered ice and common salt, both at 0° C., are mixed. The ice always has a little water attached, and this water will dissolve some of the salt to form a solution whose state must be represented by a point on the diagram (Fig. 180) between the lines BP, PD. This salt solution will practically dissolve some of the ice, and, owing to the latent heat of ice, the temperature will fall. The water formed by the liquefaction will dissolve some more salt, and so on. In this way the temperature will gradually fall, but the state of the liquid formed by the melting of the ice and the solution of the salt must always be represented by a point included between the lines BP and PD. Hence, as the temperature falls, it will finally be restricted to the single composition containing 23.8 per cent. of salt, and the temperature will be −22°, that is, the point P will be reached. No lower temperature than this can be reached, for this would involve the solidification of both the water and salt, and this operation would necessitate the evolution of heat. We thus see how it is that for every freezing mixture there is a minimum temperature, below which it is impossible to go ; so that with ice and salt, say, whether we start with the materials at 0° or at −20°, the lowest temperature produced is always −22°. It is also evident that the best results will be obtained if we take the salt and ice in the proportion of 24 parts of salt to 76 parts of ice.

The following table gives some of the common freezing mixtures :—

FREEZING MIXTURES.

Substance.	Parts.	Substance.	Parts.	Temperature.	
				At Start.	When all Snow is Melted.
Sodium chloride . .	33	Snow	100	0°	−22°
Calcium chloride (crys.)	100	"	70	0°	−54.9
Ammonium nitrate .	100	"	131	0°	−17.5
Carbon dioxide (solid) .	—	Ether	—		−77

227. Boiling-Point of Solutions.—We have in § 219 considered Dalton's laws for the vapour pressure of mixtures of liquids. If, however, we dissolve a solid body in a liquid, the vapour pressure of the solution is less than that of the pure solvent. Since a liquid boils when its vapour pressure is equal to the pressure to which it is subjected, it follows that the boiling-point of such a solution is raised. Raoult has found that for dilute solutions the lowering of the vapour pressure is proportional to the concentration, and the molecular lowering (*i.e.* the lowering produced by 1 gram-molecule of the solid dissolved in 100 grams of the solvent) is independent of the nature of the dissolved substance. As in the case of the depression of the freezing-point, the lowering in the vapour pressure has been used to determine the molecular weight of solids in solution. In determining the boiling-point of a solution, the thermometer has to be placed *in* the boiling liquid, since the temperature of the vapour given off is equal to that over the pure solvent boiling under the given pressure.

228. Thermal Phenomena accompanying Chemical Change.—We have hitherto considered the thermal phenomena which accompany physical change, and although the corresponding thermal considerations with reference to chemical change belong more especially to the science of chemistry, it will be useful here to very shortly refer to some of them.

Every chemical reaction is characterised by the evolution or absorption of a certain definite quantity of heat, so that, keeping all the external conditions the same, if the reaction takes place in the opposite sense, then the thermal phenomena simply change sign, the quantity of heat involved being the same as before. The quantity of heat involved in any given reaction depends, however, in a marked manner on the physical conditions under which the reaction takes place. Thus if 2 grams of hydrogen and 16 grams of oxygen, both in the gaseous condition, at standard pressure and temperature combine together to form water at 0°, the heat evolved by the reaction is 68834 calories. If the result of the reaction is to form steam at 100°, the heat evolved is only 58386 calories, the difference representing the heat given out by 18 grams

of steam at 100° in condensing to water at 0° (*i.e.* $536 \times 18 + 100 \times 18$). In the same way, if the result of the reaction is to give ice at 0°, the heat evolved is 70274 calories (*i.e.* $68834 + 80 \times 18$).

The above is an example of a simple reaction; as a more complicated reaction, we may take the solution of metallic zinc in dilute sulphuric acid. If 65 grams of zinc are dissolved in dilute sulphuric acid: (1) water is decomposed and two grams of hydrogen are evolved, this reaction *absorbing* 68834 calories, as in the previous example; (2) the oxygen combines with the zinc and 83500 calories are *evolved;* (3) the oxide of zinc combines with the acid forming $ZnSO_4$, and water and 23400 calories are evolved. Hence the resultant thermal effect of the whole reaction is that 38066 calories are evolved, since

$$38066 = -68834 + 83500 + 23400.$$

If one gram of diamond is converted into carbon monoxide (CO), 2140 calories are evolved; if the CO is then converted into carbon dioxide (CO_2), 5720 calories are evolved. Hence 7860 calories have been evolved in the conversion of carbon (in the form of diamond) into carbon dioxide, the reaction having taken place in two steps. If 1 gram of diamond is directly converted into carbon dioxide, the heat evolved is 7860 calories, so that the same amount of heat is evolved whether the reaction takes place in one or in two steps. This is an example of the law that when a system of bodies passes from one state to another, the quantity of heat evolved is independent of the intermediate states through which the bodies pass.

229. Curves Showing the Relations between the Temperature, Volume, and Pressure of a Body.—We have seen in the previous sections that the volume of unit mass of a substance, the pressure to which it is subjected, and the temperature bear definite relations to one another, and that if we know these three particulars we can deduce the physical state of the substance. The consideration of many properties of a substance are made much clearer by drawing curves showing the connection between these three quantities. Since, however, there are *three* of them, we cannot draw a single curve on a plane surface to represent the changes which take place in them all. A series of curves can, however, be drawn showing the relation between any two, the third being supposed to remain constant. There are three possible kinds of curves, namely, (1) those showing the relation between pressure and volume at constant temperature; (2) those showing the relation between pressure and temperature, the volume being constant; and (3) those showing the connection between temperature and volume, the pressure being constant. The first of these, in which the temperature is constant, are called *isothermals;* the second, in which the volume is constant, *isometric lines* or *isopleres;* and the third, in which the pressure is constant, *isopiestic lines* or *isobars.*

230. Isobars.—As an illustration of the use that may be made of the isobars, we will examine the form of these curves in the case of water. In order to draw the curve for a pressure of one atmosphere, we should plot the volume of 1 gram of water (in the state of ice, water, or steam, as the case may be) as ordinate against the corresponding temperature. The following table gives the volume of 1 gram of water at some temperatures:—

VOLUME OF 1 GRAM OF WATER AT ATMOSPHERIC PRESSURE.

Temperature.	State.	Volume.
$-10°$	Ice	1.0897 c.c.
$0°$	"	1.0909 "
$0°$	Water	1.0001 "
$4°$	"	1.0000 "
$+50°$	"	1.0120 "
$100°$	"	1.0431 "
$100°$	Steam	1650. "
$150°$	"	1870 "

To illustrate in a figure the isobar for the range of temperature given in the table accurately to scale would be impossible, since the change in volume in passing from water at 100° to steam at 100° is so enormously greater than any of the other changes in volume. In Fig. 181, however, some isobars are shown in a *diagrammatic* form, in particular, the increase in volume during the change into steam has been indicated as being very much smaller than it really is, and the change in volume of the solid and liquid by pressure is much exaggerated. The line *abcdef* shows the isobar for a pressure of one atmosphere. Starting at the lowest temperature, at the point *a*, as the temperature rises the ice expands. When 0° is reached the ice begins to melt, and the temperature remains constant till the whole of the ice has melted, the volume decreasing during this time, so that the part of the curve *bc* corresponds to the mixture of ice and water. From 0° to 100° the water expands, and at 100° the water begins to change into vapour, the temperature again remaining constant, while the volume changes. Hence *de* corresponds to the coexistence of liquid and vapour. When the whole of the liquid has vaporised, the temperature rises above 100°, and the isobar corresponds to a gas. Since the volume of a perfect gas varies directly as the absolute temperature, the isobar for a gas is a straight line which, if produced, would pass through the point on the diagram corresponding to the temperature $-273°$ and the volume equal to zero. At a temperature some way above the boiling-point, a vapour behaves as a perfect gas, so that the remainder of the isobar above this temperature will be a straight line passing through the point $t=-273°$ and $v=0$.

If we started with a pressure of fifteen atmospheres, at any temperature the volume of unit mass would be less than its volume under a pressure of one atmosphere, and hence the isobar for fifteen atmospheres, $a'd'e'f'$, lies below the isobar for one atmosphere. Under the increased pressure, the melting-point is lowered a little and the boiling-point is raised to 200°, and the two vertical parts of the curve, which correspond to a mixture of ice and water in the one case, and of water and steam in the other, are farther to the left and right respectively.

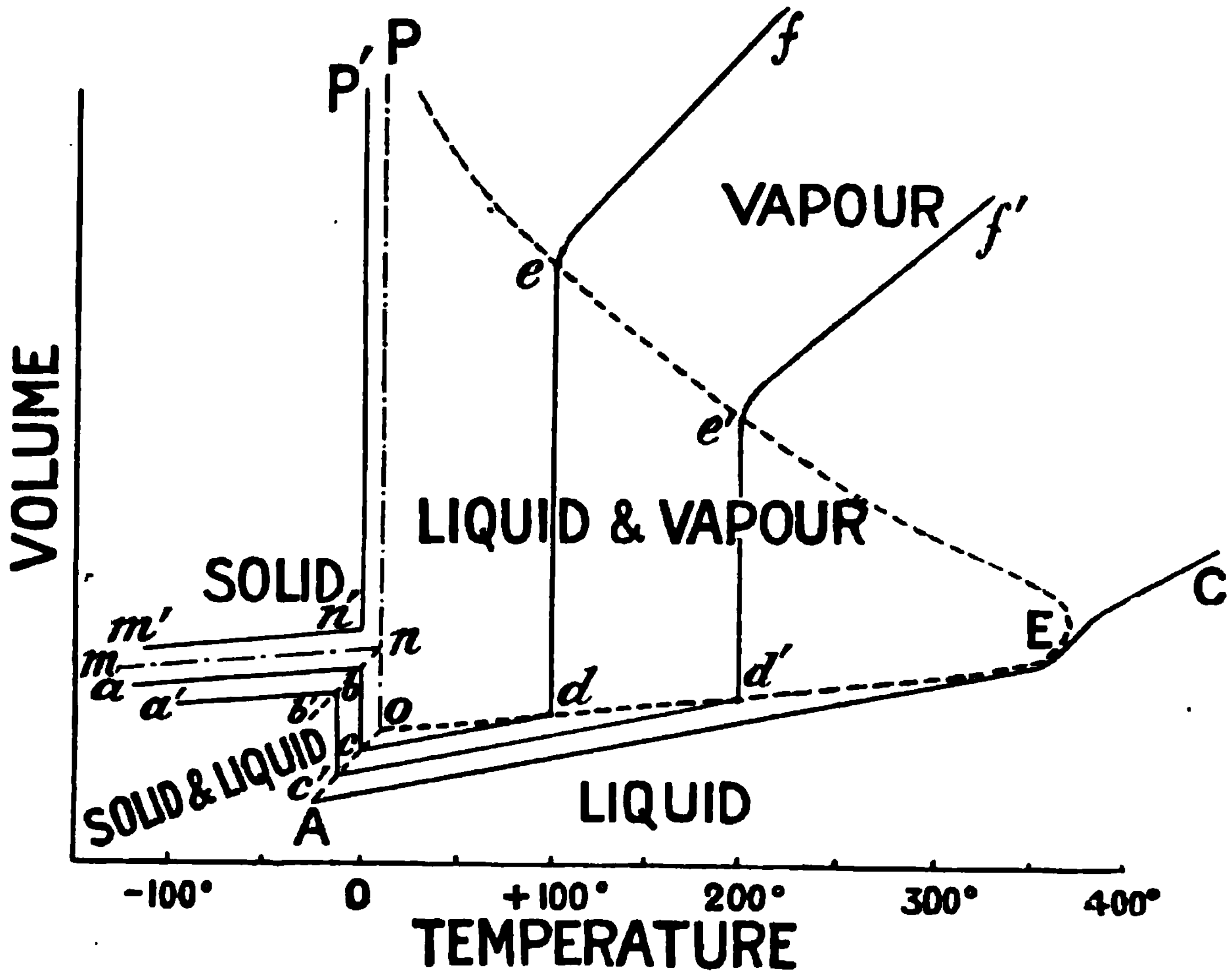


FIG. 181.

The isobar corresponding to a pressure of .04 cm. of mercury is peculiar, and is represented by the dotted lines *mn* and *oP*. Starting with the solid, at this pressure the melting-point is slightly above 0°, but at this temperature and pressure, as we have seen (§ 224), we may have all three states, solid, liquid, and vapour, existing simultaneously. Hence the isobar beyond this point is a vertical line *oP*, the part *on* corresponding to a mixture of liquid and solid, and the part *nP* to a mixture of solid and vapour. For still lower pressures there will be no liquid phase, so that the ice will sublime, and the vertical part of the isobar, *n'P'*, will correspond to a mixture of solid and vapour, the point where the whole of the solid has vaporised being off the top of the diagram.

It is thus evident that from a series of isobars for a given substance

we may draw much information as to its behaviour under various conditions of temperature and pressure, and we shall have occasion later on (§ 232) to again refer to this diagram.

231. Isothermals. — When considering the general form of the isothermal curves, it will be convenient to refer to the portion which deals with the passage from solid to liquid separately, since here we have to consider two distinct cases, namely, when the solid is denser than the liquid, as is the case with paraffin wax, and when the liquid is denser than the solid, as is the case with water. This consideration does not apply to the passage from liquid to gas, since the density of the gas is always less than that of the liquid.

Let us first take the case of a body, such as paraffin wax, in which there is expansion when the solid melts. Here, as we have seen in § 210, increase of pressure, since it increases the density of the liquid, will raise the melting-point, or, in other words, at a high pressure the liquid can pass into the solid condition at a higher temperature than it can at a low pressure. Let *ab* (Fig. 182) represent the change in volume of one gram of solid paraffin as the pressure is lowered, the temperature being kept constant and equal to t_1. At a pressure corresponding to the point *b*, let the temperature t_1 be the melting-point, so that the wax will now start melting. During the time that the solid is changing into liquid, the pressure will remain constant, while the volume increases (wax *expands* on melting), and the portion *bc* of the isothermal will be parallel to the axis of volumes, *i.e.* horizontal. When all the solid has melted, then, if the pressure is further reduced, the liquid will expand, and the line *cd* will represent the continuation of the isothermal.

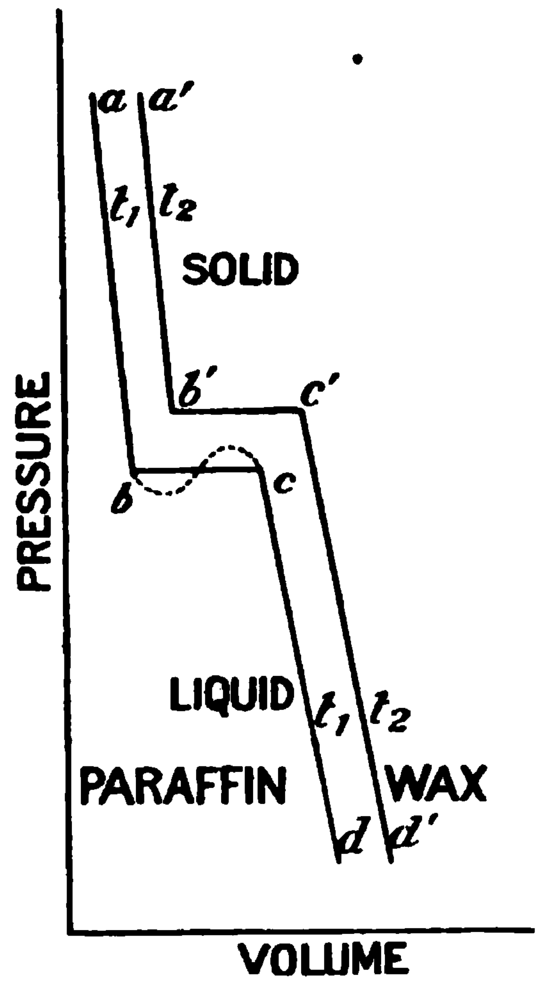


Fig. 182.

Suppose now that we start at the same pressure as before, but at a temperature t_2 higher than t_1. As the temperature is higher, the volume is greater, the pressure being the same, and the isothermal will start at the point *a′*. As the pressure is reduced, the volume will increase, and *a′b′* will represent the portion of the isothermal corresponding to the solid state. Since the temperature is now higher than before, fusion will start at a higher pressure than before, for the higher the pressure, the higher the temperature at which fusion takes place. Hence the horizontal portion *b′c′* of the isothermal, which corresponds to the coexistence of the solid and liquid states, will be at a higher pressure than before.

When all the solid has melted, the pressure will again fall and the liquid expand, so that $c'd'$ will represent the remainder of the isothermal.

Next, considering the case of a substance, such as water, in which there is expansion on solidification, so that increase of pressure lowers the melting-point, and therefore ice can only be melted by *increasing* the pressure, the temperature remaining constant. We have in this case to start with the liquid at *a* (Fig. 183), say at a pressure of a thousand atmospheres and a temperature of $-0°.76$. As the pressure is reduced the water will expand, till the pressure is reduced to 100 atmospheres, when the water will commence to freeze, and the pressure will remain constant, the volume increasing till all the water is converted into ice. When this transformation is complete, on further reducing the pressure the ice will expand. The isothermal will therefore have the form *abcd*.[1]

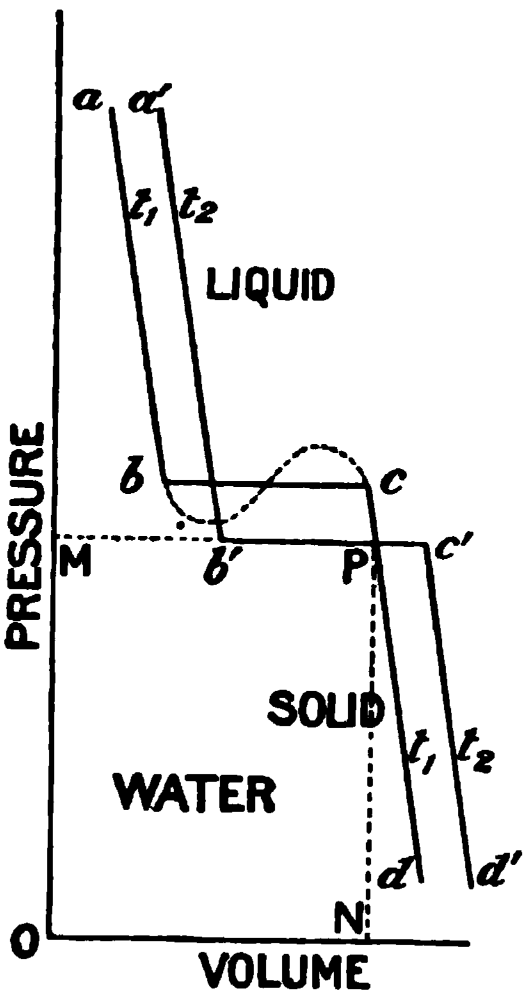


FIG. 183.

Next, suppose we start at the same pressure as before, namely, 1000 atmospheres, but at a higher temperature, say 0°. The pressure will now have to be reduced to one atmosphere before ice commences to form, so that the horizontal portion of the isothermal $b'c'$, which corresponds to the coexistence of liquid and solid, will be at a lower pressure than before, and will cut the isothermal for the lower temperature at some point P. In the first case (Fig. 182), where the body expands on fusion, the isothermals were quite distinct, and nowhere intersected. The intersection of the isothermals at P means that at the pressure OM there are two temperatures, t_1 and t_2, at which unit mass of the substance has the same volume ON. In one case, that at the temperature t_1, the substance is wholly liquid, while in the other, t_2, the substance is partly liquid and partly solid.

The general form of the isothermals for water and steam are shown diagrammatically in Fig. 184. The horizontal portions AB, CD, &c., represent the passage of the substance from the liquid to the gaseous condition, during which these two states coexist. The curve for the vapour at some distance from the point B is a rectangular hyperbola, of which the axes of pressure and volume are the asymptotes, since the vapour at

[1] If the pressure is sufficiently reduced the ice will sublime, and there will be a second horizontal portion of the isothermal corresponding to the coexistence of the solid and vapour. When the whole of the ice has sublimed, on further reducing the pressure, we should get the isothermal corresponding to a vapour.

pressures removed from its condensing point behaves like a perfect gas, and for a perfect gas Boyle's law holds, so that pv = a constant, which is the equation to such a rectangular hyperbola.

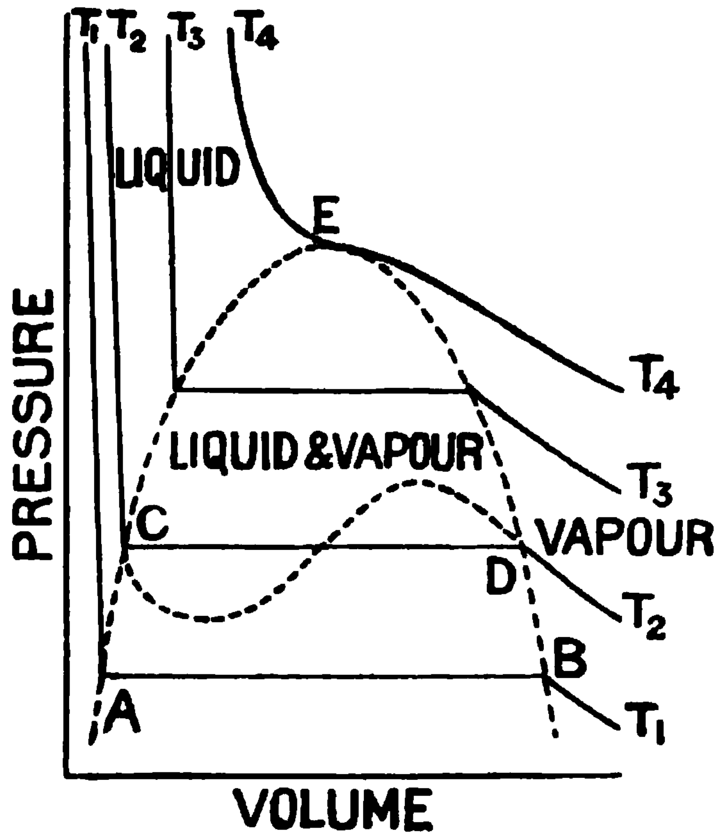


FIG. 184.

Suppose we start with an isothermal corresponding to a temperature t_1, so that vaporisation commences at a pressure represented by the point A and ends at B. Next consider an isothermal corresponding to a higher temperature t_2. In this case, since the temperature is higher than before, the vapour pressure will be greater, so that vaporisation will be able to start at a higher pressure than before. At this higher temperature and pressure, vaporisation will be complete at a smaller volume than before, so that the change in volume, represented by the length of the line CD, during the passage from liquid to gas will be smaller than before. This decrease in the change in volume during the change in state will continue as the isothermals correspond to higher and higher temperatures, till finally a temperature t_4 will be reached, such that there is no sudden change in volume, during the passage at or about the point E, from the liquid to the gaseous state. We shall in § 232 return to this subject, when considering the critical temperature of a gas.

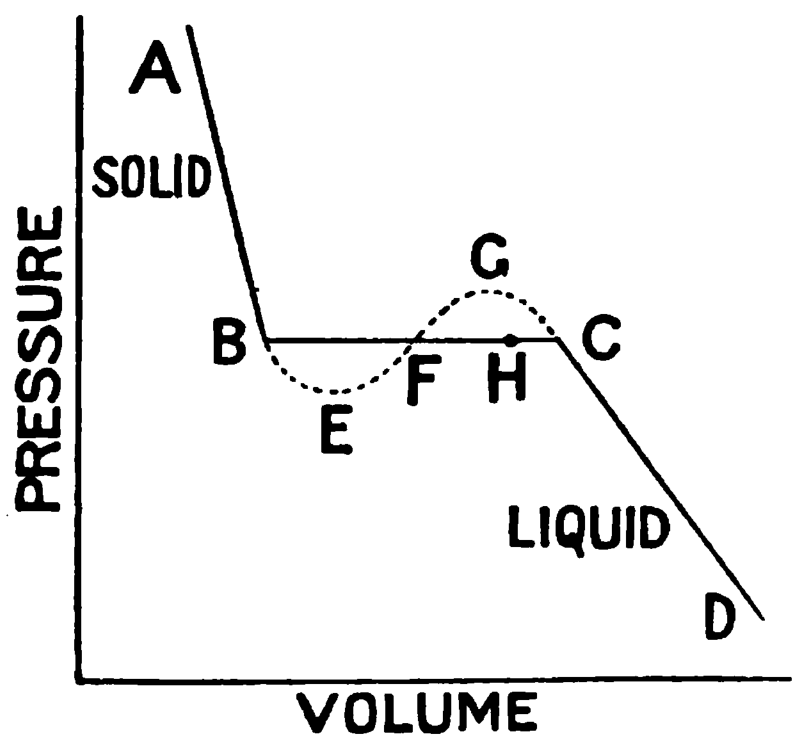


FIG. 185.

Each of the isothermals shows four points at which a sudden change of direction takes place, one each at the commencement and end of the two parts corresponding to the change of state. The broken line ABCD (Fig. 185) represents a portion of an isothermal of a body which expands on fusion, that is, the density of the liquid is less than that of the solid, the straight part corresponding to a mixture of solid and liquid. Professor James Thomson has suggested that the true form of the isothermal curve during the change of state is

as shown by the dotted line, in which no abrupt changes in the direction of the curve take place, and the substance remains homogeneous throughout. On this theory, the explanation of the fact that the observed part of the isothermal is the straight line BC is that what we observe must necessarily be the mean condition of a very large number of molecules, and thus although, if we were able to follow the change in the volume, &c., of a single molecule or small group of molecules, they would each be found to follow the dotted curve, the change from solid to liquid, or liquid to vapour, being continuous, yet in the case of a relatively large mass of liquid there are a number of such groups simultaneously going through such a change, and the integral result is represented by the straight line BC, the area of the loop BEF being equal to that of the loop FGC.

The part of the curve between E and G we cannot expect to be able to actually observe, for it represents an essentially unstable condition, as an increase of pressure is here accompanied by an increase of volume, which increase in volume itself causes a further increase of pressure, and so on. The parts BE and GC correspond, however, to stable conditions, and hence we may reasonably hope that, under special circumstances, we shall be able to obtain a body in the condition indicated by points on these portions of the isothermal.

The form of the isothermal for phosphorus is similar to that shown in Fig. 185, by the line ABCD. If, however, we take some phosphorus in the liquid condition, as indicated by the point C, it is possible, if suitable precautions are taken, keeping the pressure the same, to cool the liquid without its solidifying, so that the new condition is represented by the point H. Now the points C and H appear to be on the same isothermal, yet to pass from C to H we have lowered the temperature of the body, or, in other words, have *not* passed along an *isothermal*.

We are therefore driven to the conclusion that the points H and C are not on the same isothermal, and that the true form of the isothermal is really as shown by the dotted lines, so that the point H is on the upper part of the serpentine of an isothermal corresponding to a lower temperature than the one passing through C.

The case of a "superheated" liquid, in the same way, realises the portion BE of the isothermal corresponding to the passage from the liquid into the gaseous condition.

232. The Critical Point.—In Fig. 186 are given the isothermals for carbon dioxide, as obtained by Andrews. At low temperatures (13°.1 and 21°.5), the horizontal part of the curve, corresponding to the presence of both liquid and vapour, is very marked. The curve for a temperature of 31°.1 has, however, quite a different form, as there are no longer sharp bends in the curve. Starting with the gas at 31°.1, and increasing the pressure, the volume diminishes at first slowly, but at a pressure of about 75 atmospheres the volume diminishes rapidly, very much as happens during the condensation of a vapour, but there is no visible separation of

the carbon dioxide into two distinct conditions. As the pressure is further increased the volume diminishes, but only slowly, the rest of the isothermal curve resembling that for the liquid; and for pressures of about 90 atmospheres the volume is what we should expect it to be at this temperature, from the known coefficient of cubical expansion of *liquefied* carbon dioxide. Andrews found that at temperatures below 30°.92 there

FIG. 186.

was a clearly marked passage from the liquid to the gaseous condition, but that for all higher temperatures, starting with the gas, it was possible to compress it till it possessed the density, &c., of the liquid, but that no abrupt change from one state to the other took place. This temperature, below which the abrupt change from gas to liquid can take place, is called the *critical temperature.*

The isothermals for 32°.5 and 35°.5 show that at these temperatures there is a pressure at which the rate of change of volume is excessive, but this flattening of the curve is less marked for the higher temperature. At a temperature of 48°.1 this flattening has entirely vanished, and the curve is similar to that obtained in the case of one of the so-called permanent gases.

A similar series of phenomena are exhibited by other substances which can be obtained in both the gaseous and liquid condition. At pressures below the critical temperature, if the pressure is increased sufficiently, there is a sudden change from the gaseous to the liquid condition, while for temperatures above the critical temperature there is no such abrupt change, the substance gradually passing from the condition of a gas into that of a liquid, the two states never coexisting.

If on the diagram of the isothermals for a substance we trace a curve through the points at which liquefaction commences and ends, we obtain a curve such as is shown dotted in Fig. 184. The isothermal for the critical temperature will touch the vertex of this curve, because for all lower temperatures we get a distinct commencement and end of liquefaction. For the body in the states represented by all points included within this curve, we may have the liquid and gas existing side by side. The point E, at which the isothermal for the critical temperature touches this curve, corresponds to what is called the critical point of the substance, and the pressure and volume which correspond to E are called the critical pressure and volume of the substance respectively. It is well to remember that what we mean by the critical volume is the volume of *unit mass* of the substance at the critical temperature and pressure.

In the isobars for water, shown in Fig. 181, it will be noticed that the change in volume in the passage from the liquid to the gaseous state decreases as the pressure for which the isobar is drawn increases. The isobar corresponding to the critical pressure is a continuous curve AEC, touching the curve $odd'Ee'e$, which can be drawn enclosing that portion of the diagram corresponding to the possible coexistence of the liquid and gaseous states, at the point E. The co-ordinates of E are, of course, the critical temperature and volume respectively.

The following table contains the critical data for some substances :—

CRITICAL DATA.

	Temperature in Degrees C.	Pressure in Atmospheres.	Volume in c.c. per Gram.
Carbon dioxide	30.92	77	2.2
Sulphur dioxide	156	78.9	1.9
Ether	194.4	35.61	3.8
Water	365	195	2.3
Oxygen	−118	50.0	1.5
Nitrogen	−146	33.0	2.7
Hydrogen	−234	20.0	...
Ammonia	130	115	...
Benzene	288.5	47.9	3.3
Acetic acid	321.6	57.1	2.8

233. Density of the Saturated Vapour and of the Liquid up to the Critical Point.—At temperatures below the critical temperature, a substance may exist either as a liquid or as a saturated vapour, and hence if we plot a curve showing the connection between the density of a substance and the temperature, the pressure being always such that the liquid and vapour can simultaneously exist, we shall get a curve such as that given for carbon dioxide in Fig. 187. Thus at a temperature of 10° carbon dioxide may exist either as a liquid having a density of .85, or as

a saturated vapour having a density of .14. As the temperature rises the density of the liquid decreases, while that of the *saturated* vapour increases; and when the critical temperature is reached, the densities of the liquid and vapour are equal, so that the curves showing the density of the liquid and of the vapour meet at the point P at the critical temperature. It will be noticed, from the curve, how very rapidly the density of the liquid and of the vapour change near the critical temperature, and it will be understood why the accurate determination of the critical volume is so difficult.

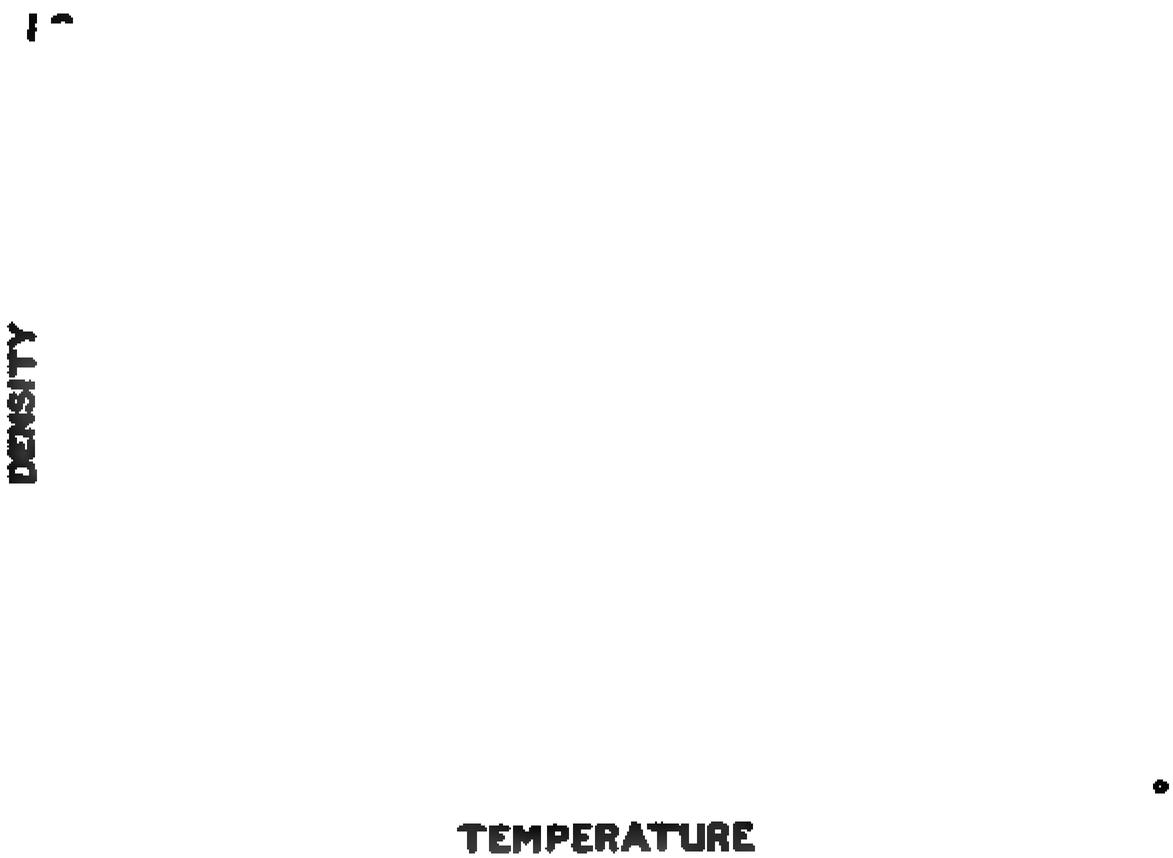


FIG. 187.

Cailletet and Mathias have, however, shown that if we take the mean of the densities of the liquid and saturated vapour at each temperature, and plot these means on the diagram, the points obtained will all lie on a straight line (QP, Fig. 187) which passes through the critical point P. Thus by drawing the density curve, and producing the diametral straight line to cut it, we obtain the density at the critical point, from which, of course, the critical volume can at once be calculated.

234. Van der Waals's Equation connecting the Pressure, Volume, and Temperature of a Fluid.—When considering the kinetic theory of gases in § 142, we showed that if V is the mean velocity of the molecules, p the pressure to which the gas is subjected,

and v the volume occupied by unit mass of the gas, then, since the density is equal to $1/v$, we have

$$\tfrac{1}{3}\overline{V}^2 = pv,$$

and that, if Boyle's law is true, it follows that $\overline{V}$ is constant at any given temperature.

If T is the absolute temperature of the gas, then by Charles's law (§ 197) $pv = RT$, where R is a constant depending on the nature of the gas. Combining this result with the one just obtained, we get

$$\tfrac{1}{3}\overline{V}^2 = RT.$$

In other words, the mean velocity of the molecules of a gas is directly proportional to the square root of the absolute temperature.

Now, in obtaining the equation $\frac{1}{3}\overline{V}^2 = pv$, we supposed that the molecules exerted no attraction or repulsion on one another between their successive impacts, and, further, we neglected the size of the molecules. Taking into account the size of the molecules and a possible attraction which the molecules might exert on one another, and which would be very similar to that we have assumed to exist in the case of liquids when dealing with capillary phenomena (§ 157), Van der Waals calculated the value of the mean velocity of the molecules, and hence deduced an equation showing the relation between p, v, and T, which should correspond to the equation $pv = RT$ that applies in the case of perfect gases. His calculations showed that the effect of the attractions between the molecules was to add a term to p, and he took it to be of the form a/v^2, where a is a constant. The effect of the finite size of the molecules was to virtually diminish the volume, v, in which the molecules can move by a constant amount b. His modified equation then took the form

$$\left(p + \frac{a}{v^2}\right)(v - b) = RT.$$

At the absolute zero, where $T = O$, it follows that $v = b$, for p and $\frac{a}{v^2}$ must both be positive, and hence $p + \frac{a}{v^2}$ cannot be zero. Thus b represents the minimum volume a gas can be made to occupy. If v is very great, *i.e.* the gas is far removed from its condensing point, the quantity a/v^2 is excessively small, and hence we get the relation that the volume of a gas diminished by the constant b (which, since by hypothesis v is large, produces little effect) is proportional to the absolute temperature. This agrees with the observed fact that a gas, when far removed from its condensing point, behaves as a perfect gas, and obeys Boyle's and Charles's laws.

By multiplying through by v^2, Van der Waals's equation may be written in the form

$$pv^3 - v^2(pb + RT) + av - ab = o.$$

Since a, b, and R are constants, if we take the pressure and temperature as having some definite values, we have a cubic equation from which to find v. It is shown, in books on the theory of equations, that a cubic equation must necessarily have either one or three real roots. Hence we must have either one or three values of v corresponding to given values of p and T. In an isothermal curve T is constant, so that a line drawn parallel to the axis of volumes, and therefore corresponding to a constant pressure, must cut the isothermal in either one place or three places. Thus the line p_1p_1' (Fig. 188) cuts the isothermal for a temperature T_1 at

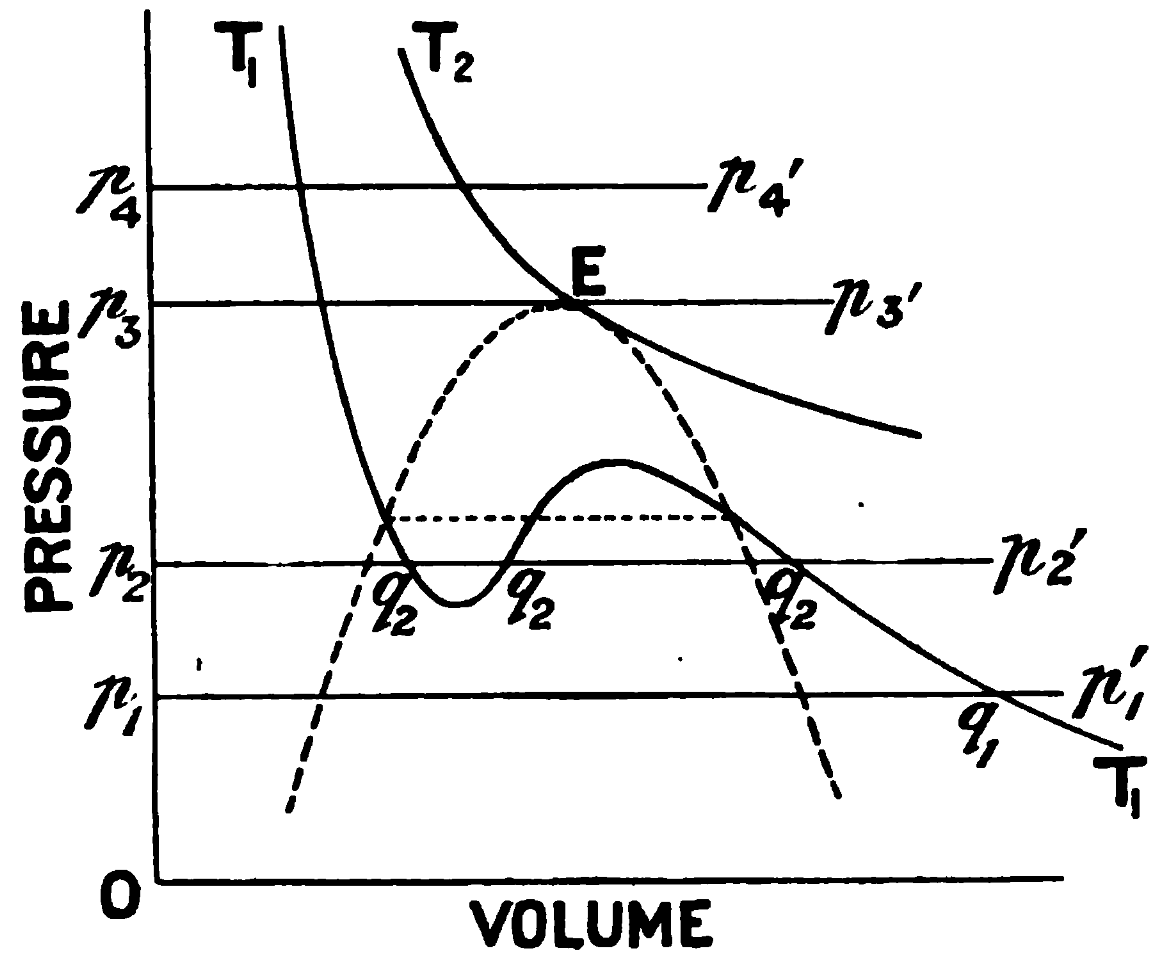


FIG. 188.

one place only, q_1. The line p_2p_2', however, if we assume the truth of James Thomson's hypothesis (§ 231), cuts the isothermal at three points q_2, q_2, and q_2. As we consider isothermals nearer and nearer to the critical point E, the three possible values for the volume get nearer and nearer, and at the critical point E they coincide. For all temperatures above the critical temperature there is only one possible value of the volume. At the critical point E the three roots of the cubic equation are all real and are equal. From this condition it can be shown, if p_c, v_c, and T_c are the critical pressure, volume, and temperature respectively, that

$$p_c = \frac{a}{27b^2}\,;\; v_c = 3b\,;\; T_c = \frac{8a}{27Rb}.$$

Hence, since the values of the constants a and b can be obtained from the deviations of a gas from Boyle's law, we can calculate the critical constants of a gas from the observations on the deviations from this law.

dioxide through the inlet on the left, the supply being regulated by the valve W, which is worked by the head B. This liquid can evaporate freely, and in doing so becomes so much cooled that the remainder solidifies. The oxygen, which is stored under pressure in a steel cylinder, enters the apparatus by the right-hand inlet, and passes up through the tube O. It then passes round the spiral S, which is immersed in the solid carbon dioxide, and thus becomes cooled to about $-70°$. Next the oxygen passes down the spiral tube D to the tube U, in the side of which there is a very small jet, which can be closed by the rod V and screw A. The compressed gas escaping by this jet and expanding becomes cooled, and this cooled gas passes up, as shown by the arrows, between the spirals of the tube D, through which the oxygen is descending, and then escapes into the air. In its passage up between the spirals, the cooled oxygen cools the spirals and the contained oxygen, so that the oxygen escaping at the jet becomes colder and colder. Each portion of oxygen as it travels down the spiral is cooled down by the escaping gas to the temperature this has acquired by its expansion at the jet, and this oxygen, when it in turn reaches the jet and expands, becomes yet further cooled. This regenerative process goes on till the escaping gas at the jet is cooled down to its liquefying point, when liquid oxygen collects in the vessel G. This vessel is of particular construction, so as to reduce the conduction of heat from surrounding objects to the liquefied gas to a minimum. It consists of a double-walled glass test-tube, the space between the walls being exhausted to the highest attainable vacuum. In such a "vacuum vessel," particularly if the outside is silvered, so as to be a very bad absorber of radiant heat (§ 246), it is possible to preserve liquid air for many hours. A third method of liquefying such gases as oxygen and air will be described in § 254.

E

FIG. 191.

If oxygen is caused to evaporate rapidly, by connecting a closed vessel containing the liquid to an exhaust-pump, such a low temperature is obtained that the air in contact with the vessel containing the boiling

liquid oxygen is liquefied at the ordinary pressure, and may be collected in a vessel placed to catch it as it drips down.

By allowing hydrogen which was cooled to -205° C., by passing first through a coil in a vessel B (Fig. 191) containing solid carbon dioxide, then through a coil in a vessel C containing liquid air, which was caused to boil rapidly by reducing the pressure, and under a pressure of 180 atmospheres to escape through the nozzle G of an apparatus somewhat similar to that shown in Fig. 190, the vessel D being itself placed in a space kept below -200°, liquid hydrogen has been found by Dewar to collect. The liquid hydrogen was thus collected in the form of a liquid even at atmospheric pressures. By introducing a glass tube filled with helium into the liquid hydrogen, a distinct drop of liquid, presumably helium, formed in the tube. Thus all the known gases have been condensed into liquids, and the term "permanent gas" has no meaning.

CHAPTER IV

CONDUCTION OF HEAT

236. Transference of Heat.—When defining the higher of two temperatures, we said that if, when two bodies are brought near each other, heat passes from the one to the other, the one from which the heat passes is said to have the higher temperature. We have now to consider the laws that govern the passage of heat from one body to another. Heat may be propagated in three ways. In the first place, heat may travel from one portion of matter to another by what is called *radiation*, and in this process the transference can take place without the intervention of matter.[1] It is by radiation that heat (and light) reach us from the sun. In the second place, heat may be propagated by the actual visible transference of matter, as in the case when a building is heated by the flow of hot water through pipes. This method of propagation is called *convection*. Thirdly, heat may be propagated by *conduction*. In this case the heat is conveyed by matter, but no visible motion of the matter itself takes place; the heat is usually considered as propagated by the warmer molecules heating the neighbouring colder molecules, and so on. Thus, when one end of a metal rod is placed in a flame and the other is placed in melting ice, it is found that heat is conducted along the rod, causing the ice to be melted.

237. Conduction.—In order to define the conductivity for heat of a body, let us suppose we had a slab of the material of thickness d, with parallel faces each of area A, and that the opposite faces are kept at the temperatures t_1 and t_2 respectively. Then heat will be conducted by the material of the slab from one face to the other. Let Q units of heat pass from one face to the other through the slab in a time τ. Then it is found that

$$Q=k\frac{A(t_2-t_1)\tau}{d},$$

where k is a constant for any one substance, independent of the thickness, area of the faces, and the difference of temperature (so long as this is not too great), but varies from one substance to another. If we make each of the quantities A, d, τ, and the difference of temperature (t_2-t_1) unity, we have that k is equal to the quantity of heat which would pass

[1] When considering the subject of light, we shall show that the energy, in the case of radiant heat, is propagated by a wave motion in the ether.

between the opposite faces of a slab of the material of unit area and of unit thickness in unit time, when the temperatures of the faces differ by unity. The quantity k is called the ***thermal conductivity*** of the substance.

The difference of temperature between the faces, $(t_2 - t_1)$, divided by the thickness, gives the change of temperature per unit length in the direction in which the heat is flowing, and is called the temperature gradient. The rate at which the temperature of a body, say a metal rod, rises when it is heated at one end, depends not only on the conductivity of the material, but also on the specific heat. Let c be the specific heat of the material and ρ its density, then the heat required to raise the temperature of unit volume through one degree is $c\rho$. Now the thermal conductivity k is the quantity of heat which would pass through a slab of the material of unit thickness and unit cross section, in unit time, when the temperatures of the two faces differ by one degree. This quantity of heat would raise the temperature of unit volume of the material through t° where t is given by the equation

$$k = c\rho . t.$$

Hence
$$t = k/c\rho.$$

The quotient $k/c\rho$, or the coefficient of conductivity divided by the heat required to raise the temperature of unit volume through one degree, is called the *diffusivity* or coefficient of *thermometric conductivity* of the material.

238. The Measurement of the Conductivity of Solids.—If one end of a long bar is heated, and a series of thermometers are placed in small holes drilled in the bar, the readings of the thermometers will increase. The thermometer nearest the heated end will rise first, the others following in succession. After a time the temperature of all parts of the bar will become constant, but diminishing gradually from the heated end to the other end. When this occurs, the heat supplied to the bar at the hot end during each second is exactly equal to that lost by radiation and conduction from the sides and the cold end. Let a curve AB (Fig. 192) be drawn such that the abscissæ represent distances along the bar, measured from the heated end, and the ordinates represent the corresponding temperatures. If we consider two cross sections of the bar at M and N, the temperatures at these points being represented by MR and NS, the difference in temperature between these two sections is equal to $\overline{MR} - \overline{NS}$ or to $\overline{RP}$. Of the heat which crosses the section of the bar at M, part is conducted on and crosses the section at N, while the rest is radiated from the outside surface of the bar between the two sections. By taking the distance $\overline{MN}$ between the two sections sufficiently small, the proportion of heat lost by radiation from the edges of this small section of the bar bears so small a proportion to the heat conducted through the section, that we may neglect it. Also, since the points R

and S are so near together, we may take the temperature curve between these points as being a straight line tangential to the curve. Hence if A is the area of cross section of the bar, and k the conductivity of the

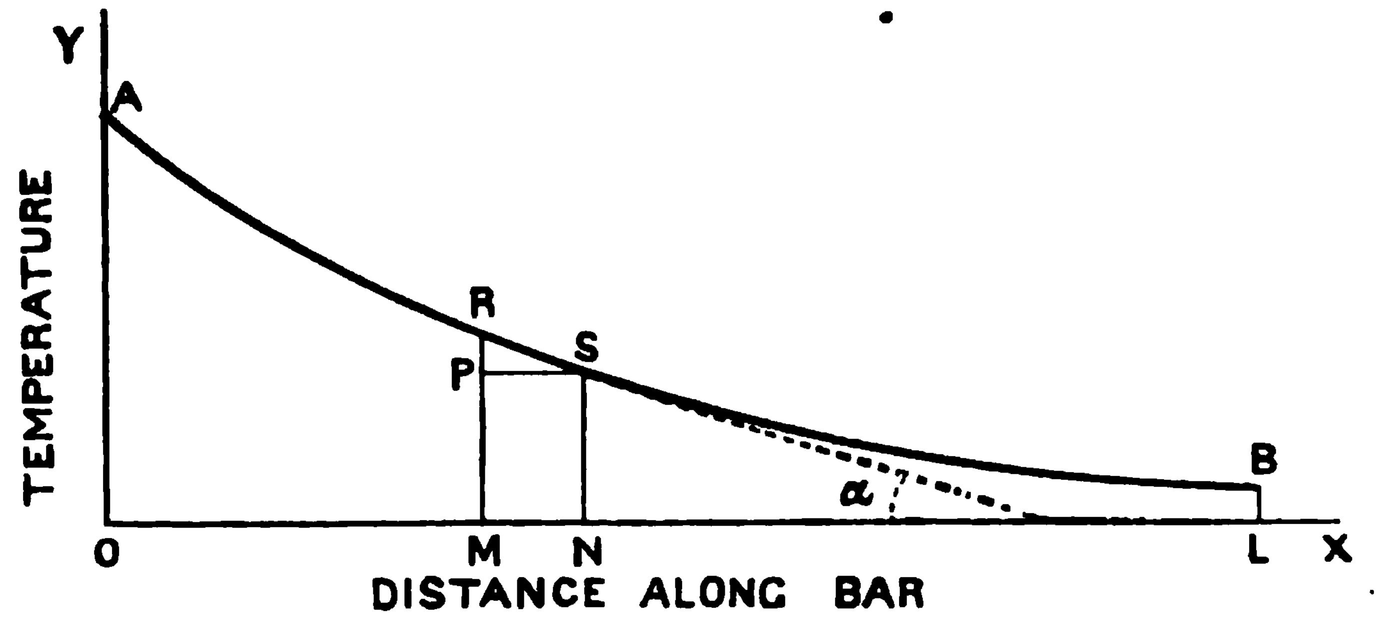


FIG. 192.

material of which the bar is composed, we have that the quantity of heat Q which passes through the section MN in unit time is given by

$$Q=kA\frac{\overline{PR}}{\overline{MN}}.$$

If the tangent to the temperature curve at R or S makes an angle α with the axis of X, then, since $\overline{MN}=\overline{PS}$, we have

$$\frac{\overline{PR}}{\overline{MN}}=\tan\alpha.$$

Hence $$Q=kA\tan\alpha.$$

If the bar is sufficiently long, there will be some point L which is at the same temperature as the surrounding air, then all the heat which passes through the section MN must be lost by radiation from the surface of the bar between N and L. Hence if we can measure the heat lost by radiation by the portion $\overline{NL}$ of the bar, we shall know Q; and then from the cross section, A, of the bar and the angle, α, made by the tangent to the temperature curve at N, with the axis, we can calculate the conductivity k.

In order to determine the heat lost by radiation, a separate experiment is made, in which a short bar of the same material, and having its surface in the same condition as that of the long bar, is heated uniformly to a temperature slightly higher than that at the point N, and is then allowed to cool by radiation, the temperature being read at short intervals by means of a thermometer placed in a small recess in the bar. Knowing the specific heat of the material and the mass of this short bar, the quantity of heat lost in one second by unit length of the bar at dif-

ferent temperatures can be obtained in the same way as was done in the case of the calorimeter in § 201. Hence the quantity of heat lost by each unit of length of the first bar between N and S per unit time at its temperature, as given by the temperature curve, can be obtained, and the sum of these quantities of heat gives *Q*.

Experimenting by this method, Forbes obtained 0.207 for the conductivity of wrought iron at 0°, and 0.157 at 100°.

In order to compare the relative conductivities of bars, use is made of the relation that the conductivities are proportional to the squares of the distances of points of equal temperature from the source of heat. In order to make the loss of heat from the surface of the bars the same in all cases, they are either coated with lamp-black or electroplated with silver. In the experiments of Wiedemann and Franz, one end of the rods were heated to 100° in a steam bath, and the temperature at different points was obtained by means of a thermo-electric junction.

Another method of measuring the conductivity of fairly good conductors of heat, such as metals which can be obtained in large pieces, is illustrated diagrammatically in Fig. 193. The material to be tested is in the form of a thick block A. One face of this block closes the circular

FIG. 193.

end of a cylindrical steam chamber C. This chamber is surrounded by a larger chamber B, and they are both supplied with drain tubes, by means of which the condensed steam may be drawn off. The opposite face of A closes two exactly similar cylindrical boxes, which are filled with pounded ice. Two fine holes are bored into A in planes parallel to the faces to which the cylinders are attached, and into these are inserted

two delicate thermometers, or thermo-elements (§ 498), which serve to measure the difference in temperature between two planes in the material at a distance d apart. The outside cylinder B preserves the inside cylinder from loss of heat by radiation and conduction, except on the face in contact with A. Hence, if w is the weight of water collected at C in a second, this means that wL thermal units have been conducted into A, through an area A, where L is the latent heat of steam, and A is the area enclosed by C. In the same way, if w_1 is the weight of water collected at F per second, and L_1 is the latent heat of ice, w_1L_1 units of heat have been given up by an area A of the surface of the substance A. Now if the outside cylinders B and D are sufficiently large, so as to include a wide anulus of the surface of A, the flow of heat within the block will, as indicated by the dotted lines, be quite uniform between the portions of the opposite surfaces enclosed by C and E. Thus all the heat that enters from C and leaves to E will travel across an area A in the planes containing the thermometers T_1 and T_2. Thus

$$Q = wL = w_1L_1 = \frac{kA(T_2 - T_1)}{d},$$

so that k can be calculated.

In the case of bad conductors of heat a somewhat similar method is employed, only, since these have to be taken in thin slabs, it is possible so to arrange matters that the loss of heat by the edge is very small compared to the quantity of heat which passes through, for in a thin disc the area of the cylindrical surface bears only a small ratio to the area of the faces.

A diagrammatical section of an arrangement used by Lees is shown in Fig. 194. Two thin discs A, A′ of the material are each placed between two copper discs B, B′ and C, C′, while the discs B, B′ are fixed to the opposite sides of a flat coil D of insulated platinoid wire. An electric current is passed through this coil, through the wires E, E′, and the amount of heat liberated per second is calculated from the value of the current and the resistance of the coil (§ 493). A very little glycerine is placed between the discs A, A′ and the copper discs so as to insure good thermal contact. The copper discs are such good conductors that the temperatures as given by the thermo-elements T_1, T_2, T_3, T_4, which fit into fine holes drilled in these discs, may be taken as being the temperatures of the surfaces of the discs A, A′. Thus, knowing the quantity

Fig. 194.

of heat supplied to D, the temperatures of the faces of the discs A, A′ and their thicknesses and areas, we can calculate the thermal conductivity as before. A correction can, if necessary, be applied for the loss of heat by the edges of the discs, and conduction along the wires E, E′ and the thermo-element wires.

The following table gives the thermal conductivity of some solids at ordinary temperatures :—

THERMAL CONDUCTIVITY IN CALORIES/CM.SEC.

Silver	1.096	Marble	0.005
Copper	1.041	Glass	0.0025
Aluminium	0.344	Cork	0.0007
Iron	0.167	Sulphur	0.00067
Zinc	0.303	Paraffin	0.0002
Granite	0.005	Horn	0.00009

239. Temperature of the Earth's Crust.—An interesting problem in connection with the conduction of heat in solids is furnished by the crust of the earth. The surface of the earth is alternately heated and cooled, and thus we have a series of waves of heat due to the heating during the day and to the cooling during the night, as well as a series due to the heating during the summer being above the average, and that in the winter below the average, which start at the surface and travel towards the centre of the earth. The diurnal wave is only sensible at depths of 2 or 3 feet, while the annual wave can be traced to a depth of about 50 feet.

DIURNAL VARIATIONS MAY 1898. MONTREAL.

FIG. 195.

The kind of change obtained at moderate depths by Callendar is shown in Fig. 195. The three curves show the temperature recorded by thermometers placed at depths of 4 inches, 10 inches, and 20 inches respectively, below the surface of the ground at Montreal, Canada. The

ground consisted of a light brown sandy soil with turf on the surface. It will be noticed that at 20 inches below the surface the diurnal change can hardly be detected.

Fig. 196.

In Fig. 196 the mean temperature of the air and of the earth at depths of 4 inches, 20 inches, and 9 feet, are given for a whole year.

While the maximum air temperature occurs in August, at a depth of 9 feet the maximum does not occur till early in October.

The *mean* temperature of the earth's crust is, however, found to increase steadily as we descend. The rate of increase with depth varies very much with the geological conditions, but amounts on the average to about 1° C. for a depth of 28 metres or 30 yards. Since heat always flows from places of high to places of low temperature, this increase of temperature with the depth shows that there must be a continuous flow of heat from the interior of the earth to the surface, and from the conductivity of the crust and the temperature gradient the loss of heat in a year can be calculated. From the present rate of loss we can then calculate what the temperature of the earth must have been in times past, and in this way Lord Kelvin has shown that it cannot be more than 200,000,000 years since the earth was a molten mass on the outside of which a solid crust was just forming.

240. The Measurement of the Conductivity of Liquids.—When measuring the conductivity of solids, we are not troubled with convection; in the case of fluids, however, it is extremely difficult to arrange for conduction to take place unaccompanied by convection. When the lower strata of a fluid are heated, the fluid expands and becomes less dense, and hence the heated portions stream up through the colder, which sink to the bottom. These convection currents tend to equalise the temperature throughout the fluid mass. When a liquid is heated at the top, convection currents are to a great extent eliminated, and the passage of heat to the lower strata is excessively slow, except in the case of mercury and molten metals.

The conductivity of water has been measured by Bottomley, using a modification of the form of the experiment originally due to Despretz. The water is contained in a cylindrical vessel A (Fig. 197), and the heat is supplied by gently pouring a stratum of hot water on a wooden float placed on the surface of the water. Four thermometers, T_1, T_2, T_3, T_4, are placed with their bulbs in the positions shown. The two thermometers T_1 and T_2 give the difference of temperature of the faces of a horizontal stratum of the liquid of known thickness. The quantity of heat which flows through this stratum per second is obtained by observing the change of temperature of the mass of the liquid below the stratum, the thermometer T_3, the bulb of which extends from the stratum to nearly the bottom of the vessel, giving the mean temperature. As soon as the wave of heat reaches the bottom of this thermometer, as indicated by the thermometer T_4, the experiment is stopped. In this way Bottomley found the number 0.002 for the conductivity of water.

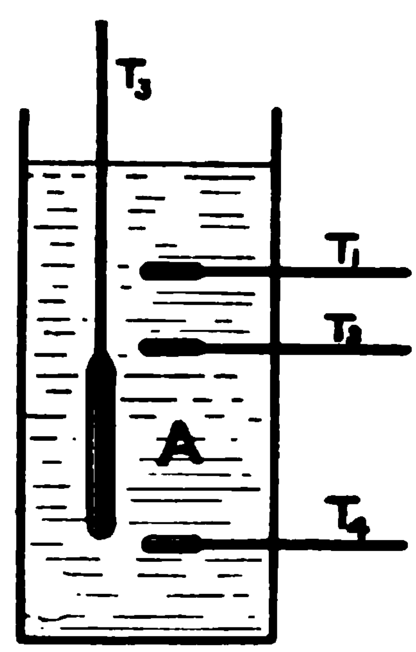


FIG. 197.

Lees has employed the disc method described in the preceding section to measure the conductivity of liquids. The liquid is enclosed by an ebonite ring fixed between the discs A and C (Fig. 194). A correction is applied for the conductivity of the ring, and the discs are placed horizontal, with the hot side of the liquid film uppermost, to reduce the effects of convection currents.

The following table gives the thermal conductivities of some liquids:—

THERMAL CONDUCTIVITY OF LIQUIDS IN CALORIES/CM.SEC.

Water	0.0014	Ether	0.0003
Glycerine	0.0007	Mercury	0.0152
Ethyl alcohol	0.0004		

241. The Measurement of the Conductivity of Gases.—The determination of the conductivity of a gas is a problem of even more difficulty than the measurement of the corresponding quantity in the case of a liquid, for convection currents play even a more prominent part than before, and it is difficult to separate the effect due to conduction and radiation.

It follows from the kinetic theory of gases, and has been found by experiment, that the conductivity of a gas is independent of the pressure so long that this is not reduced so much as to make the mean free path (§ 141) of the molecules of appreciable magnitude with reference to the dimensions of the vessel enclosing the gas.

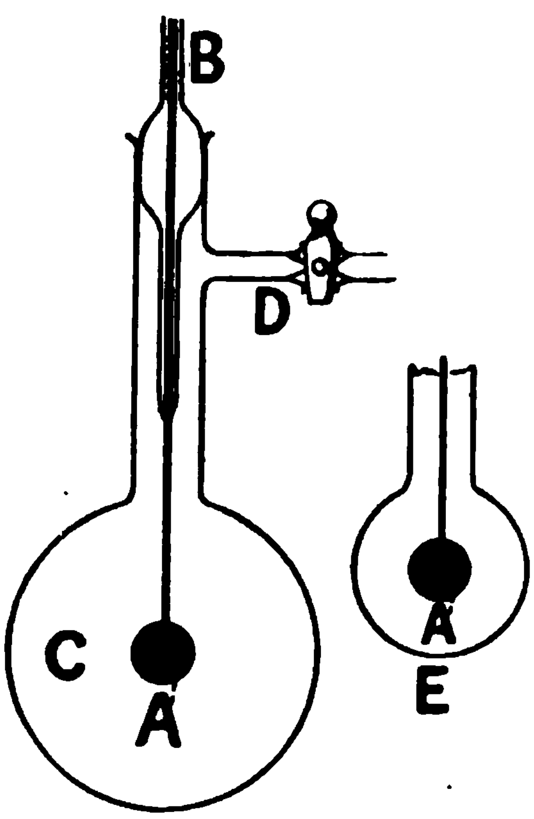


FIG. 198.

In the experiments made by Müller, using a method previously employed by Kundt and Warburg, a small spherical glass globe A (Fig. 198), containing mercury, is used as the hot body. This globe forms the bulb of a mercurial thermometer, on the stem of which there is an enlargement B that acts as a stopper to the spherical glass vessel C. The air from between A and C can be exhausted through a side-tube D. The whole apparatus having been heated to a known temperature, say 100°, and the air exhausted to the required amount, the globe C is plunged into a vessel containing water which is kept at a known constant temperature, or into a mixture of ice and water. The times taken for the thermometer to fall through a given interval, say 10°, are then noted. Knowing the "water equivalent" (§ 201) of the thermometer bulb A, the quantity of heat which must be lost for the temperature to

fall 10° can be calculated, and hence the quantity of heat which passes from A to C, through the air which is enclosed between them, is known. This loss of heat may be due to three causes—(1) conduction, (2) radiation, and (3) convection.

Of these the third, namely, that due to convection, varies with the pressure. Now it is found experimentally that, in the case of air, the rate of loss of heat by A, at any given temperature, decreases as the pressure is decreased till a pressure of about 15 cm. of mercury is reached, and then remains constant down to a pressure of about 0.1 cm. Hence for pressures between 15 cm. and 0.1 cm. the loss of heat due to convection currents is inappreciable, and therefore the observed rate of loss is due to conduction and radiation only. In order to separate the effect of these two causes, two methods have been employed. In one of these the globe is exhausted to the best vacuum obtainable, particular attention being paid to the removal of the last traces of mercury vapour from the globe. The rate of loss of heat is then measured, and since there is now practically no gas present, the loss is taken to represent the loss due to radiation only. Deducting this loss from that obtained when the pressure was about 2 cm., the loss due to conduction alone is obtained. As an example, the times taken to cool from 59°.88 to 58°.88 at different pressures were as follows :—

Pressure	1.5 cm.	1.0 cm.	0.5 cm.	0.0 cm.
Time to cool 1°. . .	3.6 sec.	3.7 sec.	3.6 sec.	8.0 sec.

Thus at pressures of 1.5, 1.0, and 0.5 cm. the rate of cooling was constant, while at the best attainable vacuum it was only about half as much.

Another method of allowing for the radiation is to repeat the observations with an outside vessel, such as E, of a different size to the first. Since the thickness of air between the hot body and the walls of the vessel is different in the two cases, the loss of heat by conduction will be different. The loss of heat by radiation will, however, at any given temperature, be the same as before. Hence, by making experiments with outer vessels of two sizes, the loss of heat by radiation can be allowed for.

The conductivity of air obtained by Müller was 0.000056 $\dfrac{\text{calories}}{\text{cm.sec.}}$ Kundt and Warburg found the conductivity of hydrogen to be 7.1 times that of air, and that of carbon dioxide 0.59 times. Combining these results with the value for air given above, we have :—

Conductivity of Heat in Gases.

Hydrogen	0.00040 calories/cm.sec.
Air	0.000056 „
Carbon dioxide	0.000033 „

242. The Spheroidal State.—If a metal plate is heated to a temperature very considerably above 100° C., and a few drops of water are thrown on it, these do not immediately boil away, as occurs when the temperature of the plate is only a few degrees above 100°. The general appearance of the drops of water resembles that of mercury on a glass dish, for if the drops are small they are almost spherical,[1] while as the quantity of water is increased they become more and more flattened. That evaporation is going on all the time is shown by the gradual diminution in the size of the drops. The liquid in this experiment is said to be in the spheroidal state, and the slow evaporation is due to the fact that it is not in contact with the hot metal, but is separated from it by a thin layer of vapour, which is being continually renewed by evaporation from the lower surface of the drop. The vapour being, like all gases, a very bad conductor of heat, the water only slowly acquires heat from the metal, while the heat lost, owing to the evaporation which is taking place at the under side, due to the latent heat of vaporisation, is sufficient to keep the temperature of the drop below the boiling-point.

If the drop is placed on a flat and level metal plate, it is possible to see between the drop and the surface of the plate, thus showing that there is no true contact between the liquid and metal. The phenomenon is also exhibited by other liquids, the only condition being that the metal plate must have a temperature considerably higher than the ordinary boiling-point of the liquid.

[1] The spherical shape is due to the action of surface tension (§ 157).

CHAPTER V

RADIANT HEAT

243. Provost's Theory of Exchanges.—As we have seen in the last section but one, when a hot body is suspended in a gas, cooling takes place due to two distinct causes. In the first place there is loss of heat due to convection currents set up in the gas, and to conduction through the gas, both of which depend on the presence of matter, while in the second place heat is radiated in all directions, independently of the presence or absence of matter. Hence the rate of cooling depends on two distinct terms, one due to convection and conduction, and the other to radiation. In the case of convection, the heat energy is conveyed by the circulation of comparatively large groups of molecules at the higher temperature, all moving together, while in conduction the energy is handed on from molecule to molecule, owing to their impacts. In radiation the energy is conveyed by waves set up in the luminiferous ether by the vibrations of the heated molecules, or rather atoms within the molecules. While in the case of convection and conduction the medium through which the heat is propagated becomes heated, this is not necessarily the case with radiation. Thus the radiant heat which reaches the earth from the sun does not heat the intervening interplanetary space. If the body is suspended in a vacuum, then we have only to deal with radiation. Suppose, therefore, that we have a vessel which is free from air, and the walls of which are kept at a constant temperature by the vessel being immersed in a water bath, and that we introduce into this vessel a body at a higher temperature than the walls of the vessel. Under these circumstances the hot body will lose heat by radiation, and this passage of heat from the hot body to the walls will continue till its temperature becomes the same as that of the walls. Again, if a body at a lower temperature be introduced, heat will be radiated by the walls to the body, and its temperature will rise till equality of temperature is reached. If, now, the body is removed from the enclosure and placed in an enclosure at a lower temperature, its temperature will immediately begin to fall, and it will commence to lose heat by radiation. Since, therefore, there can be no property in the walls of the new enclosure which would enable them, when there is no material connection between them and the body, to cause the body to start radiating, we conclude that the body is radiating all the time, but that when its temperature remains constant it is gaining just as much energy,

due to the radiation from the walls of the enclosure, as it is itself radiating. Arguing in this way, Provost propounded what is known as Provost's theory of exchanges. According to this theory, all bodies are always radiating heat, and the reason the temperature of the body in the above examples becomes stationary is that, when the temperatures of the body and of the enclosure are equal, the amount of heat radiated by the body is just equal to the amount it receives, due to the radiation of the walls of the enclosure.

244. Instruments for Measuring Radiant Heat.—When radiant heat of any wave-length (§ 265) is absorbed by a body, the energy of the radiation is converted into heat, and the temperature of the body rises. Hence since, as we shall see, we can deduce, from the rise in temperature of the body, the energy which must have been converted into heat to produce this rise in temperature, if it were possible to prepare a body which was a perfect absorber of radiation of all wave-lengths, the measurement of its rise in temperature would give the energy of the incident radiation. Although no such perfectly black body is known, yet a surface coated with lamp-black or platinum-black absorbs such a large proportion of the incident radiation of all wave-lengths as to sufficiently nearly fulfil the conditions for practical purposes. The various instruments which have been devised for measuring radiation differ from one another in the way in which the rise in temperature of the blackened body, on which the radiation falls, is measured.

The oldest arrangement is the thermopile, which consists of a number of bars, alternately of antimony and bismuth, soldered together in series. The bars are so arranged that the alternate junctions between the two metals are near together, as shown diagrammatically in Fig. 199, some insulating material, such as mica, separating the adjacent bars. The junctions of the bars are coated with lamp-black, so that when radiation is allowed to fall on the end of the thermopile it is absorbed, and thus the temperature of the junctions is raised. Now when the alternate junctions of such an arrangement of metals are heated, and the ends of the series are connected to a galvanometer G, an electric current is produced, which causes the galvanometer to be deflected. The magnitude of the current, and therefore also of the galvanometer deflection, depends on the amount by which the temperature of the face of the thermopile on which the radiation falls exceeds that of the other face. Hence the galvanometer deflection is a measure of the rise of temperature of the face, and also of the energy of the incident radiation. As the mass of metal in an ordinary thermopile is very considerable, the rise in temperature produced by the conversion of a small quantity of radiation into heat is very small, and so the arrangement is not very sensitive, and is unsuited for

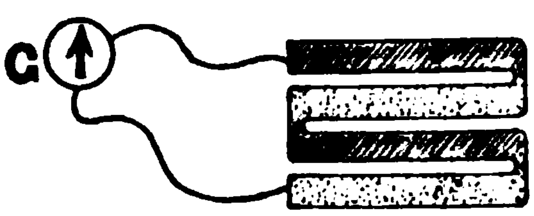


FIG. 199.

measuring small quantities of radiation. A modified form of thermopile, in which the thermopile and galvanometer are combined in a single instrument, has been designed by Boys, and is called a radio-micrometer. The thermopile proper consists of two excessively small bars of antimony and bismuth, which are soldered to the edge of a plate of blackened copper about 2 mm. square, which receives the radiation. The other ends of the antimony and bismuth bars are soldered to a light copper loop, which is suspended by a very fine quartz fibre between the poles of a powerful magnet. When the copper plate is heated by the radiation, a thermo-electric current of electricity is set up in the copper circuit, which therefore tends to turn and set itself at right angles to the lines of force of the magnet (§ 510). Thus from the deflection of the circuit, as shown by a beam of light reflected from an attached mirror, the rise in temperature of the copper, and hence the amount of the incident radiation, can be determined.

An instrument in which the rise in temperature of a blackened metal strip is measured in another way is the bolometer invented by Langley. The thin strip of blackened metal forms one of the arms of a Wheatstone's bridge (§ 488), and when the temperature rises the resistance of this strip increases, so that if before the rise in temperature the galvanometer was undeflected, this increase of resistance of one arm will upset the balance of the bridge, and the galvanometer will be deflected. The intensity of the radiation which falls on the strip will then be proportional to the galvanometer deflection, at any rate if the deflection is not very great.

245*. Equality of the Emissive and Absorptive Powers of a Body.—By the absorbing power of a body is meant the fraction of the incident radiation which a body is able to absorb. It has been found by experiment, and we shall deal with this subject later (§ 246), that a surface coated with lamp-black is capable of absorbing practically all the heat energy which falls on it, and hence its absorptive power is unity.[1] A body of which the absorptive power is unity is often called a "*perfectly black*" body. The incident energy which is not absorbed by a body is, if we confine our attention to opaque bodies, reflected. Thus if a quantity E of radiant heat energy is incident on a surface of which the absorptive power is a, the quantity of energy absorbed is aE, and the quantity reflected is $E(1-a)$.

By the emissive power of a body is meant the ratio of the quantity of heat energy emitted by one square centimetre of its surface, under given conditions, to the quantity emitted under the same conditions by one square centimetre of a body of which the absorptive power is unity (practically, lamp-black is taken as the standard).

Suppose that we have a body A, of which the absorptive power is

[1] Since a body cannot absorb more heat than is incident on it, no body can possess an absorptive power greater than unity.

unity, so that its emissive power is also unity, enclosed in an enclosure, and that the absorptive power of the walls of this enclosure is a, the emissive power being e, and the area of its surface S. Let the temperature become the same throughout, so that, according to Provost's theory, the quantity of heat radiated by the body A is equal to the heat received from the enclosure. Let A lose E units of heat by radiation per second. Then the walls of the enclosure will absorb Ea units, and reflect $E(1-a)$ units per second. This heat reflected from the enclosure will be completely absorbed by the body A, since its absorptive power is unity. Hence in unit time the body A will lose a quantity of heat $E-E(1-a)$ or Ea; and since its temperature remains constant, this means that the enclosure must radiate Ea units which are absorbed by A. Now since the area of the enclosure is S, the radiation per square centimetre is Ea/S. If, however, the enclosure had been perfectly black, it would have absorbed E and radiated E. Hence the heat emitted per square centimetre would have been E/S. Now the emissive power e is the ratio of the actual emission to the emission of a similarly situated perfectly black surface. Thus

$$e=\frac{Ea}{S}\Big/\frac{E}{S}=a.$$

In other words, the emissive power is equal to the absorptive power. Although we have proved this relation as a deduction from Provost's theory of exchanges, we shall in the following section describe the methods by which it can be proved experimentally.

246*. Measurement of the Coefficients of Absorption and Emission.—In the preceding section we have considered how the coefficients of absorption and emission of radiant heat are defined, and we now proceed to consider how these quantities are measured experimentally. Since the character of the radiation emitted by a body is often quite different from the character of the radiation absorbed—thus lamp-black absorbs light but only emits invisible heat rays, at any rate at ordinary temperatures, some method of measuring the intensity of the radiation has to be adopted which takes account of the radiation whatever the wave-length. The only way of doing this is to measure the radiation by the rise in temperature it will produce in a body which absorbs all kinds of radiation *equally* well, so that we in this way measure the total energy corresponding to the radiation of all wave-lengths. Of course it would be better to employ a body which would absorb the *whole* of the radiation, but no such body exists. Lamp-black, although it does not absorb quite the whole of the incident radiation, yet possesses the property of absorbing very nearly, if not quite, the same proportion of the incident radiation whatever the wave-length, and so this substance is taken as a standard.

In a series of experiments on the absorbing power of opaque bodies,

Provostaye and Desains used a thermometer bulb, which was coated in turn with the various substances, to measure the radiation absorbed. The incident radiation was concentrated on the thermometer bulb by means of a lens fixed in the side of the closed box in which the thermometer bulb was fixed. If Q is the quantity of radiation (measured in ergs) which falls on the thermometer bulb in unit time, and A is the absorption coefficient for the whole bulb when coated with a given substance, the heat absorbed in unit time is QA. The thermometer will rise in temperature till the heat lost by radiation is exactly equal to the heat absorbed, and after this the temperature, t, will remain constant. If w is the water value of the thermometer bulb and its contents, and v the fall of temperature in one second, or the velocity of cooling, as it is called, at a temperature t, the heat lost by the bulb in unit time is wv. Hence

$$QA = wv.$$

Next, if the bulb is coated with lamp-black, the intensity of the incident radiation being the same as before, and if it comes to a stationary temperature t_1, the velocity of cooling being now v_1, we have, as before—

$$QA_1 = wv_1.$$

Hence
$$\frac{A}{A_1} = \frac{v}{v_1}.$$

Now the values of the quantities v and v_1 can be measured by noting the time taken by the thermometer to cool through a given temperature interval at the temperatures t and t_1 respectively, and so the absorbing power of the medium, compared to that of lamp-black taken as unity, can be calculated.

In the experiments two sources of radiation were used, namely, the sun and an Argand burner, and the results obtained are shown in the following table, the absorption of lamp-black being taken as unity :—

COEFFICIENTS OF ABSORPTION.

	Sun.	Argand Burner.
Lamp-black	1.00	1.00
Platinum-black	1.00	1.00
White-lead	0.09	0.21
Cinnabar	...	0.28
Silver in powder	...	0.21
Gold in leaf	0.13	0.04
Silver in leaf.	0.07	...

In order to measure the emissive power of various substances, the faces of a hollow metal cube are coated with the different substances, one face being coated with lamp-black, and the amount of radiation emitted, when the cube is filled with water at 100° C., is measured by means of

a thermopile placed at a constant distance from the cube. In performing the experiment, great care has to be taken to screen the thermopile from all radiation except that coming from the face of the cube which is being measured. It has been shown that many measurements made were completely vitiated by the fact that the screens, used to confine the beam of incident radiation, reflected part of the radiation they received back to the radiating surface where reflection again took place, and so some of this doubly reflected heat was thrown on to the thermopile. This error was particularly marked when dealing with the emission of polished metallic surfaces where the emission is small and the reflecting power is considerable.

In the following table is given the value of the emissive power of some bodies in terms of that of lamp-black taken as unity :—

COEFFICIENTS OF EMISSION.

Lamp-black . . .	1.00	Gold-leaf . . .	0.04
Frosted silver . .	0.05	Burnished platinum .	0.10
Burnished silver . .	0.02	Copper-leaf . . .	0.05

In order to prove directly, by experiment, that the coefficients of absorption and emission for any given substance are equal, the apparatus shown in Fig. 200 has been devised by Ritchie. The two hollow metal drums B and C are filled with air, and are connected by a glass tube which is partly filled with some liquid, such as sulphuric acid. The drum A is also hollow, and can be filled with hot water. The faces of B and A, turned towards the right, are coated with lamp-black, and the faces of A and C, turned towards the left, are coated with silver-foil. The position of the liquid column having been noted when the whole instrument is at the same temperature, hot water is placed in A, and it is found that the liquid column does not move, showing that the drums B and C are receiving the same amount of heat from A. Now the drum B receives the heat emitted by a silver surface, the heat being absorbed by lamp-black, while the heat received by C is emitted by a lamp-black surface and absorbed by a silver surface. The heat received being the same, it shows that although the quantity of heat emitted by the silver surface is small, yet the lamp-black absorbing all this heat, the result is the same as when

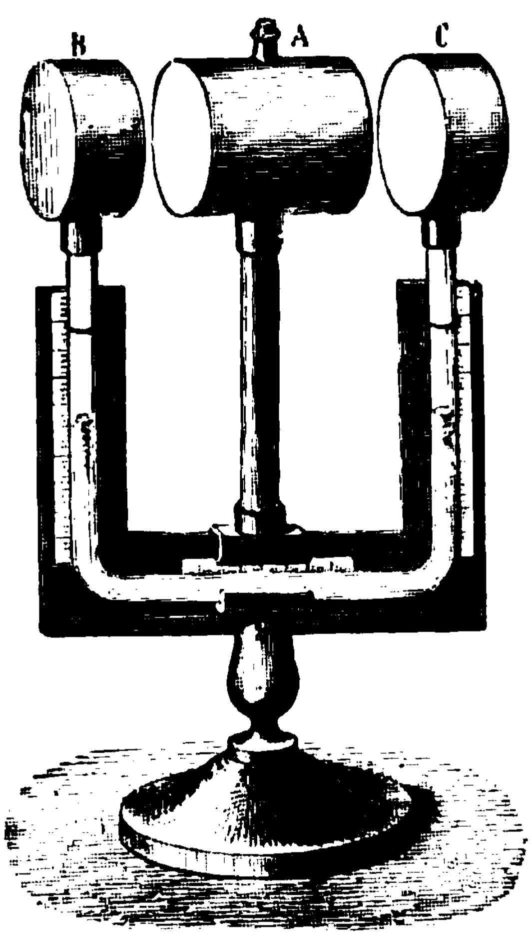


FIG. 200.
(From Ganot's "Physics.")

the large amount of heat radiated by the lamp-black surface falls on the silver surface, for in this case only a small proportion of the incident heat is absorbed.

247*. The Relation between the Amount of the Radiation and the Temperature of the Body.—In order to determine the connection between the amount of the radiation and the temperature of a body, Dulong and Petit used as radiating body the spherical bulb of a large thermometer. The bulb was heated to about 300°, and then introduced into a hollow copper sphere, which was kept at a constant temperature, and the sphere was exhausted. The temperature of the bulb was indicated by the position of the mercury thread in the stem of the thermometer, which projected from the copper vessel, and was read at equal short intervals of time.

The fall of temperature v in one second they called the velocity of cooling, so that if W is the water value of the thermometer bulb, the loss of heat in unit time, Q, is given by $Q = Wv$. As a result of their experiments, Dulong and Petit came to the conclusion that if t is the temperature of the radiating body, and t' that of the chamber, then

$$v = k(a^t - a^{t'}),$$

where k and a are constants depending on the nature and area of the surface of the radiating body. Dulong and Petit's law, which is quite empirical, has, however, been found only to hold over a small range, and Stefan, from an examination of their results, has been led to the conclusion that the total radiation emitted by a body is proportional to the fourth power of the absolute temperature. Thus if T_1 is the absolute temperature of the body, the total radiation will be represented by aT_1^4, where a is a constant depending on the extent and nature of the surface of the body. In the same way the heat radiated by the walls, if they are at an absolute temperature T_0, will be proportional to T_0^4. The quantity of this radiation absorbed by the body will be aT_0^4, since the emissive power and absorbing power of a body are the same. Hence the total loss of heat, Q, by the body in unit time is

$$Q = a(T_1^4 - T_0^4).$$

If S is the area of the radiating surface of the body, then $a = Sc$, where c is a constant depending on the nature of the surface only. Hence

$$Q = Sc(T_1^4 - T_0^4).$$

If the temperature of the enclosure is the absolute zero, and that of the body 1°, so that $T = 0$ and $T_1 = 1$, and the surface of the body is unity, we get $Q = c$, or the quantity c represents the heat radiated per second from a square centimetre of the surface of the body, when the temperature of the body is 1° on the absolute scale, and the enclosure is at the absolute zero.

If the difference between the temperature of the body and the enclosure is θ, we have

$$Q=Sc\{(T_0+\theta)^4-T_0^4\},$$

and if θ is small, so that we may neglect terms in θ^2 and higher powers. Then

$$Q=4ScT_0^3\theta.$$

Hence for small changes in the difference of temperature θ between the body and the enclosure Q is proportional to θ.

In the case of a body surrounded by a gas, as we have already pointed out, the cooling is partly due to convection currents and conduction. In such a case Newton supposed that the rate of cooling, *i.e.* the quantity of heat lost by the body in unit time, was proportional to the difference in temperature between the body and the surrounding medium. This law, which is known as Newton's law of cooling, only holds good for small excesses of temperature. For such excesses, however, as ordinarily occur in calorimetry Newton's law is sufficiently accurate.

248*. Measurement of Specific Heat by the Method of Cooling.—According to Newton's law of cooling, the quantity of heat Q^1 lost by a body during the time t, when its temperature is θ degrees above the surrounding medium, is given by

$$Q^1=kS\theta t,$$

where S is the area of the cooling surface, and k is a constant dependent on the nature of the surface.

If in a time t the temperature falls by an amount $\delta\theta$, the quantity of heat lost must be $Ms\delta\theta$, where M is the mass of the body and s is its specific heat. Hence

$$Ms\,.\,\theta\,.\,\delta\theta\,.=Q^1=kS\theta t.$$

If now the experiment be repeated, using the same radiating surface and starting at the same temperature θ, and the time t^1 be noted in which a second body of mass M^1 and specific heat s^1 cools through $\delta\theta$, we shall have

$$M^1s^1.\delta\theta=kSt^1,$$

and therefore

$$\frac{Ms}{M^1s^1}=\frac{t}{t^1}.$$

Hence, if we know M, M^1, t, and t^1, we can obtain the ratio of the specific heats of the bodies.

In an actual experiment the bodies to be experimented on are contained in a calorimeter, the outer surface of which is coated with lamp-black. This calorimeter is suspended inside a vessel with double walls, the space between the walls being filled with water so as to keep the temperature of the enclosure constant. Of course, due allowance must

be made for the water value of the calorimeter, thermometer, and stirrer. This method of measuring specific heats is found only to work satisfactorily in the case of liquids, since it is only with these that the contents of the calorimeter can be kept at a uniform temperature throughout during the cooling, this condition being obtained by continuous stirring.

The further consideration of radiant heat will be deferred till the chapters dealing with the emission, absorption, &c., of light, since there is no sharp physical line of demarcation between what we recognise by one set of senses as radiant heat, and what we recognise by our sense of sight as light.

CHAPTER VI

THE MECHANICAL THEORY OF HEAT

249. Theories as to the Nature of Heat.—Up to the end of the eighteenth century there existed two rival theories as to the nature of heat. According to one of these theories, known as the caloric theory, heat was supposed to be a subtle, elastic, imponderable fluid called caloric, which permeated all kinds of matter existing in the interstices between the molecules. According to the other theory, which was only held by very few, heat was supposed to be due to the rapid motion of the molecules of matter.

It was well known that heat could be produced by friction or percussion, and the supporters of the caloric theory explained these facts by supposing that in the case of percussion the caloric was squeezed out of the body, and hence flowed into a neighbouring body such as a thermometer, and, in the case of friction, that during the friction some of the body was rubbed off, and that the capacity of matter for caloric was less in the form of a powder than in the form of a solid block. That this explanation of the production of heat by friction was untenable was first shown by Count Rumford in 1798.

Being struck by the large amount of heat developed when cannon were being bored at the arsenal at Munich, Rumford performed an experiment in which a blunt steel borer was rotated while kept pressed against the bottom of a hole in a large mass of gun-metal. The borer was rotated nearly a thousand times, and the heat developed was sufficient to raise the temperature of the whole block, which weighed 113 lbs., about 70° F., while the amount of metal rubbed off from the bottom of the hole was only 837 grains troy. Rumford, in the account of his experiments, draws attention to the fact that the supply of heat obtained in this way from a given lump of metal seems quite inexhaustible, and hence cannot be a material substance, but must be "motion."

The supporters of the caloric theory, however, still maintained that the source of heat was the abraded metal, till this explanation was completely refuted by an experiment performed by Davy. He rubbed together two blocks of ice at a temperature below 0° C., and found that heat was developed and the ice melted. Since it was allowed by the calorists that water contained more caloric than ice, if we can produce water by the friction of ice, the heat developed must be due to some other cause than the extrusion of caloric.

We have seen, when dealing with radiant heat, that a hot body is continually radiating heat into surrounding space; and when we come to the consideration of the subject of light, we shall see that there is conclusive evidence that radiant heat, after it leaves the hot body, exists as a wave-motion in some medium surrounding the body. Now in order to set up waves in a medium, we must have a body which is itself in motion in the medium. Thus, since a hot body can set up such waves, we infer that it must be in a state of motion. Also, since the highest-power microscope is quite unable to detect any motion in a hot body, we infer that this motion must be a motion of the molecules as a whole, or of the parts of a molecule, or both combined.

We are hence reduced to the theory that heat is a "mode of motion."

250. Dynamical Equivalent of Heat—First Law of Thermo-Dynamics.—In Rumford's experiments, the heat produced in the cannon was indirectly due to the work done by the horse which turned the boring tool, and it is obviously of interest to see what connection there is between the work done by the horse and the amount of heat produced. We shall see in later sections, that whenever mechanical work is converted into heat, or mechanical work performed at the expense of heat, there exists a *constant* relation between the work done and the heat produced or lost. The quantity of work which must be done in order that, when all the work is converted into heat, the unit quantity of heat energy may be produced is called the *mechanical or dynamical equivalent of heat.* If J is the value of the mechanical equivalent, then the relation between the work W converted into heat and the quantity of heat H produced is given by the equation

$$W=JH.$$

This equation, which we shall justify subsequently, expresses symbolically what is known as the first law of thermo-dynamics, which may be stated as follows :—*Whenever mechanical energy is converted into heat, or heat into mechanical energy, the ratio of the mechanical energy to the heat is constant.*

251. The Determination of the Mechanical Equivalent of Heat.—The first to experimentally show that the first law of thermodynamics is true, and determine the value of the mechanical equivalent of heat, was Joule, who between 1840 and 1878 carried on a classic series of experiments on this subject, in which he showed that the value for the mechanical equivalent was always the same, although the methods employed for converting the mechanical energy into heat differed greatly.

The first method employed by Joule consisted in measuring the heat developed when a known amount of work was done in stirring water. The apparatus used consisted of a copper vessel B (Fig. 201), inside which a brass paddle-wheel worked. A system of partitions were fixed

within the vessel, so that the vanes of the paddle could just pass, the object of these partitions being to prevent the water as a whole assuming a motion of rotation. The paddle was rotated by means of two weights E and F, which were attached to strings wound round the axle A of the paddle, which was so arranged that the weights could be wound up without turning the paddle. The rise in temperature of the calorimeter and its contents caused by allowing the weights to fall twenty times was obtained, and knowing the water value of the calorimeter and contents, the number of heat units produced could be calculated. The work done is the product of the sum of the weights of E and F into the total height through which they fall. Corrections have, however, to be applied for the fact that when the weights reach the floor they are moving with a finite velocity v, and that their kinetic energy is destroyed in the impact.

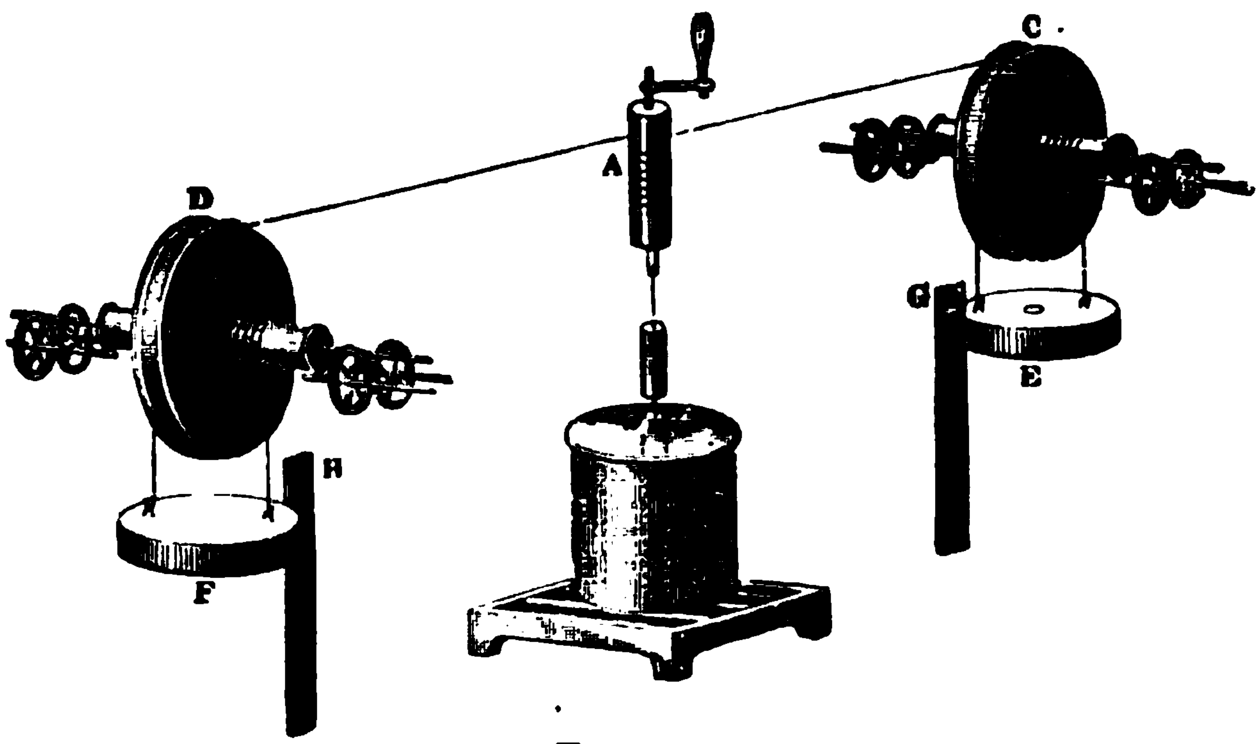


FIG. 201.
(From Ganot's "Physics.")

The height through which a body, falling freely, would acquire a velocity v has therefore to be deducted from the actual fall. Another correction has to be applied, to allow for the effect of the elasticity of the strings on which the weights are hung, for this causes the paddle to rotate a little after the weights have reached the ground. The weights themselves have in addition to be reduced by the weight, which, when the two strings are detached from the axle A and joined together, added to E or F, will just cause the weights to move uniformly. This weight represents the allowance to be made on account of the friction of the pulleys D and C and the rigidity of the string. Lastly, a correction was made for the fact that some of the mechanical energy was spent in the production of sound, and the magnitude of this correction was roughly obtained by noting the work which had to be done to make the string of a violoncello produce a sound that could be heard at the same distance as was that produced by the instrument during the fall of the weights.

In addition, Joule made a series of experiments in which the water was replaced by mercury, and also one using the friction of one iron ring against another, both being immersed in mercury.

The numbers obtained for the value of the mechanical equivalent were practically the same in all cases. Joule expressed his results in terms of the mercury-in-glass thermometer, but they have been reduced to the air thermometer by Rowland, and give the value of the energy which must be converted into heat to raise the temperature of one gram of water from $14°.5$ to $15°.5$ as

$$4.182 \times 10^7 \text{ ergs.}$$

Rowland has made some very careful measurements of the mechanical equivalent of heat by the method of stirring water, and has employed a method of measuring the mechanical work done, which was also used by Joule in his later experiments. Since this method has considerable advantages over that described above, it is worth while describing it. A diagrammatic plan of the arrangement is shown in Fig. 202. The calorimeter, like Joule's original one, had a paddle-wheel BB, and there were fixed vanes CC to prevent the water being set in rotation. The paddle-wheel was driven by means of a pulley A and a belt EF, but the calorimeter, instead of being fixed, was suspended by means of a fine wire, so that it was free to rotate about a vertical axis coinciding with that about which the paddle-wheel turned.

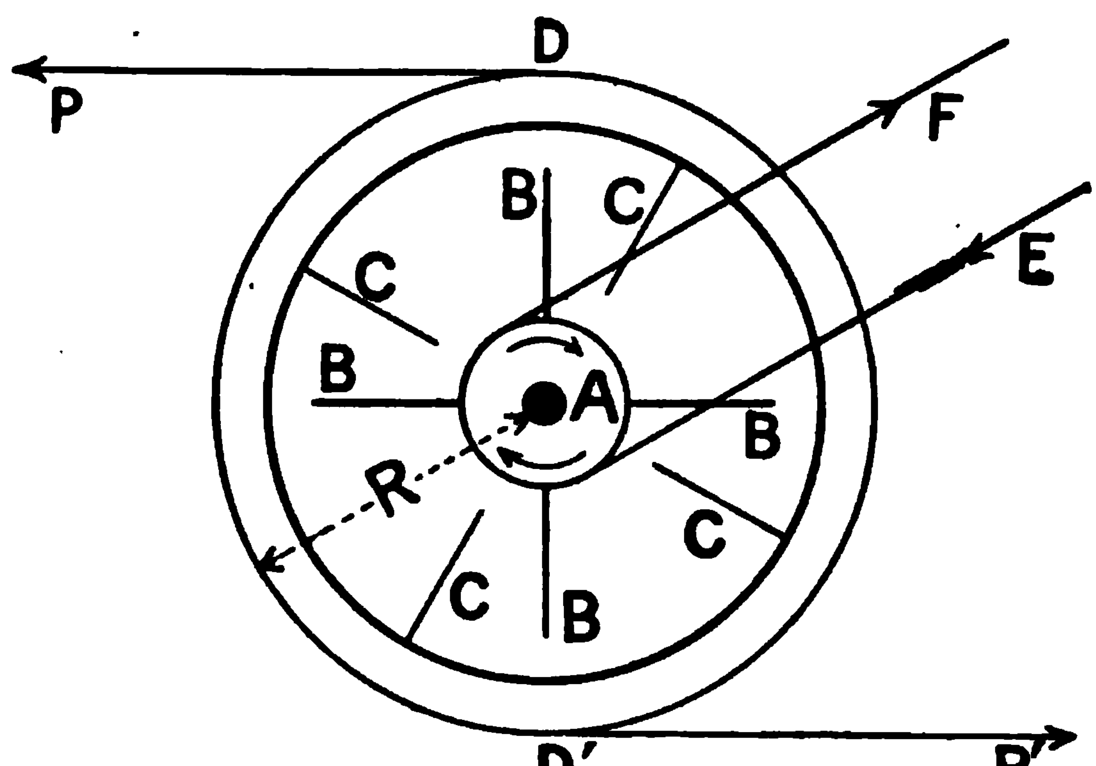


FIG. 202.

Owing to the viscosity of the water, the calorimeter tends to rotate in the same direction as the paddle, and to prevent this, two strings, DP, D′P′, were attached to the circumference of a disc which was itself fixed to the calorimeter, and these strings were pulled with a force p just sufficient to keep the calorimeter from rotating when the paddle was turning at a uniform speed. If the radius of the disc is R, the couple due to the two parallel forces p acting along DP and D′P′ is

$$2Rp.$$

Now as action and reaction are equal and opposite, and as the couple

acting on the calorimeter due to the rotation of the paddle is $2Rp$, this also must be the couple which resists the motion of the paddle. The heat generated in the calorimeter is therefore that produced by the paddle working against a resisting couple $2Rp$.

Now, the work done against this resisting couple during one complete rotation of the paddle is $2\pi \cdot 2Rp$. For the turning moment of a couple $2Rp$ is the same as that of a force $2Rp$ at the end of a lever of unit length. Suppose now we had a cylinder of which the radius was unity, and on this cylinder wound a string, the end of the string being pulled with a force $2Rp$. During one whole turn of the cylinder a length of rope $2\pi \times 1$ would be wound up, and the point of application of the force $2Rp$ would be moved through the same distance, so that the work done would be $2\pi \cdot 2Rp$. But the rope produces a turning moment of $2Rp$ on the cylinder, so that the work done when overcoming this turning moment for one whole turn is $4\pi Rp$.

If the paddle makes n revolutions, the work W done during this time is

$$W = 4\pi n Rp.$$

If the water value of the calorimeter and its contents is C, and the rise in temperature during n revolutions, corrected for radiation in the manner described in § 201, is t, then

$$J = \frac{4\pi n Rp}{Ct}.$$

The tension of the strings DP and D′P′ was supplied by passing these strings over pulleys, and attaching weights. If the *sum* of the two weights is w, the tension in each string is $wg/2$, and this is equal to p. The number of turns was determined by means of a counter attached to the spindle which carried the paddle-wheel.

Rowland obtained, as a result of his experiments, the value 4.1899×10^7 ergs for the value of J, in terms of the calorie at 15°. By measuring the heat generated by an electric current in a platinum wire, and a knowledge of the electrical energy spent, Griffiths has obtained the value 4.1940×10^7 ergs in terms of the calorie at 15° (see § 494).

A change in the unit of mass alters not only the unit of heat, but also, and in the same proportion, the unit of work, for we measure the heat in terms of the amount required to raise the temperature of unit mass of a standard substance through a given range, and the unit of energy is that possessed by unit mass when moving with a velocity of $\sqrt{2}$[1] times the unit velocity. Hence a change in the unit of mass does not affect the value of the mechanical equivalent.

A change in the unit of length, since it affects the unit of energy but not the thermal unit, will affect the value of the mechanical equivalent, as will also obviously a change in the temperature scale.

[1] Since kinetic energy $= \frac{1}{2}mv^2$, if $m = 1$, we have, when $v = \sqrt{2}$, the kinetic energy is unity.

A change in the unit of force, *i.e.* changing from the absolute system to the gravitational system, will change the value of the mechanical equivalent.

In the following table, the value of the mean of the numbers obtained by Joule, Rowland, and Griffiths is expressed in various units :—

MECHANICAL EQUIVALENT OF HEAT.

Unit of Temperature.	Temperature at which the Thermal Unit is defined.	Unit of Time.	Unit of Mass.	Unit of Length.	Unit of Force.	J.
1° C.	15° C.	sec.	gram	cm.	Dyne	4.189×10^7
1° C.	15° C.	sec.	gram	cm.	Weight of a gram at lat. of Greenwich	42690
1° F.	59° F.	sec.	pound	foot	Poundal	2504.7
1° F.	59° F.	sec.	pound	foot	Weight of a pound at lat. of Greenwich	778.1
1° C.	15° C.	sec.	pound	foot	Weight of a pound at lat. of Greenwich	1400.6

For many purposes we shall find it convenient to measure quantities of heat not in calories, but in ergs, the relation between the two being W (ergs) $= JH$ (calories).

252. Work done by a Gas during Expansion at Constant Pressure.—Suppose we have m grams of a gas enclosed within a cylinder, having a cross section A, by means of an air-tight, weightless, and frictionless piston B (Fig. 203), and that the pressure acting on the upper side is p dynes per square centimetre. When the temperature on the absolute scale of the gas is T, let the piston be at B. Next, let the gas be heated at constant pressure to a temperature T_2, the piston being driven back to B′. The total force acting upon the upper side of the piston is pA, and the piston has been driven back against this force for a distance BB′ by the expanding gas. Hence the work done by the gas in expanding against the external pressure p is $pA \times \overline{BB'}$. If the distance between the piston and the bottom of the cylinder at the temperatures T_1 and T_2 is h_1 and h_2 respectively, then the original volume of the gas is h_1A, and the final volume is h_2A. Hence the increase in volume is $A(h_2 - h_1)$, or

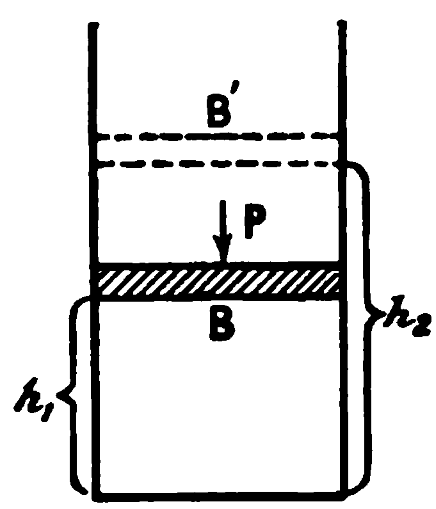


FIG. 203.

$A \cdot \overline{BB}$. Calling this increase of volume v', the external work done by the expanding gas is pv'; that is, is equal to the product of the pressure into the *increase* in volume.

253. Calculation of the Value of the Mechanical Equivalent from the difference between the Specific Heats of a Gas.—Using the notation of the last section, the quantity of heat which has to be supplied to the gas to raise its temperature from T_1 to T_2 at constant pressure is $m(T_2 - T_1)C_p$, where C_p is the specific heat at constant pressure. If, now, the gas is heated through the same range, but is not allowed to expand, the piston being fixed in the position B, the heat which has to be supplied is $m(T_2 - T_1)C_v$, C_v being the specific heat at constant volume.

The molecules of the gas have been heated through the same range of temperature in the two cases, but in the first case an amount of *external* work $v'p$ has been performed, while in the second case no external work has been done. If the molecules of the gas exert an attraction one on another, some work will have been done in the first case in separating them, since, as the gas has expanded, the mean distance between the molecules has increased. As we are unacquainted with the law governing the attraction between the molecules, we cannot calculate this work done against molecular attraction, but we may for the present indicate it by the symbol $f(v', T)$, this being chosen to remind one that it is probably dependent on the increase of volume and the temperature.

We have, therefore, that in the case of the expansion at constant pressure, in addition to the heat spent in warming the molecules, an amount of external work pv' done against the external pressure, and an amount of internal work $f(v', T)$ done against the attraction of the molecules. In the case of heating the gas at constant volume, no external work is done, and also, since the mean distance between the molecules remains unchanged, no internal work is done, so that the heat employed $m(T_2 - T_1)C_v$ is used exclusively in raising the temperature of the molecules.

It follows from the above that

$$J\{m(T_2 - T_1)C_p - m(T_2 - T_1)C_v\} = (v'p + f(v', T),$$

where J is the value of the mechanical equivalent. In the case of a perfect gas, where the molecules exert no attraction on one another $f(v', T)$ is zero. Also, from Charles's law, if v is the original volume of the gas, we have

$$\frac{v+v'}{v} = \frac{T_2}{T_1},$$

or, if ρ is the density of the gas at the temperature T_1 and under a pressure p, $v = \frac{m}{\rho}$, and

$$v' = \frac{m(T_2 - T_1)}{\rho T_1}.$$

Hence, for a perfect gas,

$$J(C_p - C_v) = \frac{p}{\rho T_1}.$$

Although air is not a perfect gas, yet Joule has shown that the term $f(v', T)$ is very small, hence, neglecting it, we can from the known values of C_p, C_v, ρ, m, and T_1, calculate J.

For air, the value of C_p is 0.238, while the ratio $C_p/C_v = 1.41$, hence $C_v = 0.169$. At a pressure of 76 cm. of mercury, *i.e.* 1013300 dynes per square centimetre, and at a temperature of 0° (*i.e.* $T_1 = 273°$) the density of air is 0.001293. Hence

$$J = \frac{p}{(C_p - C_v)\rho T_1} = \frac{1013300}{.069 \times .001293 \times 273}$$
$$= 4.16 \times 10^7.$$

This method of calculating the mechanical equivalent was first used by Mayer, who assumed, as we have done, that no heat is employed in doing internal work in the case of air, although at that time no experiments had been made to test this point. This question was first investigated experimentally by Joule, by means of the experiments described in the following section.

254. Internal Work done when a Gas Expands.—If the molecules of a gas exert an attraction one on another, then, when the gas expands, work must be done in increasing the mean distance between them. Hence if a gas is allowed to expand in such a way that it does no external work and its temperature falls, this will show that internal work has been done which has necessitated the consumption of a certain quantity of heat, so that the loss of this heat has lowered the temperature of the gas. On the other hand, if no such alteration of temperature takes place we may infer that there is no internal work done on expansion, and hence that the molecules do not exert any appreciable force on each other.

In order to allow a gas to expand without doing external work, Joule allowed the gas to expand from a vessel, A, in which it was compressed to about 22 atmospheres, into another, B, which was exhausted. The receivers were both immersed in the same water bath, which was kept well stirred, and of which the temperature was indicated by a very sensitive thermometer. The vessel A became cooled, since the gas when rushing out acquired kinetic energy; this kinetic energy was, however, entirely destroyed in B, and hence on this account just as much heat was liberated in B as was absorbed in A, and the temperature of the water bath which contained both vessels would ot alter. At the end of the experiment the volume of the gas is twice as great as at the commencement, but no external work has been done against the atmospheric pressure, since this pressure has not been driven back. If, then, any

change in temperature takes place, it must be due to internal work done in separating the molecules of the gas. Joule could detect no such change in temperature, and he concluded that no internal work was done. It must be remarked, however, that since the mass of the water and of the containing vessels was very great compared with that of the gas, with this form of apparatus he would hardly have been able to detect a change of 2° C. in the temperature of the gas.

Another series of experiments, by a method which was capable of indicating small changes of temperature, was therefore conducted by Joule in conjunction with Lord Kelvin.

The principle on which this method depends may be explained as follows :—

Let the original pressure and volume of unit mass of a gas be p and v, and the final pressure and volume p' and v'. Further, let the passage of the gas from one state to the other be made by means of the arrangement shown in Fig. 204, in which A and B are two pistons connected by a rod which passes air-tight through a partition separating the spaces C and D, and which move without friction in two cylinders, the cross sections s and s' of these cylinders being in the ratio of v to v'. If the spaces C and D between the two pistons are vacuous, the work done by the gas on A while it moves through unit distance to the right is ps, while the work done by B in pushing the gas forward is $p's'$. Hence the difference between these two quantities of work is $ps - p's'$, or, since $s/s' = v/v'$, this difference is proportional to $pv - pv'$. Now if Boyle's law holds for the gas $pv = p'v'$, hence on the whole no work is done on or by the double piston. Next let the space C be filled with gas at the pressure p, and the space D with gas at the pressure p', and let these two spaces be connected by a tube, E, in which is a diaphragm, F, pierced with a very small hole. The gas will gradually pass through this hole, and, as is evident, if the double piston is moved so as to keep the pressure in C constant and equal to p, the pressure in D will also be constant and equal to p'. When each piston has passed through unit distance, a certain mass of the gas will have passed from C to D, its pressure changing in the process from p to p'. The gas escaping into D has done no work in forcing the piston B back, since the pressure of the gas acting on A will, as we have seen, exactly do the requisite work. This energy is of course supplied by the pump used to keep the pressure to

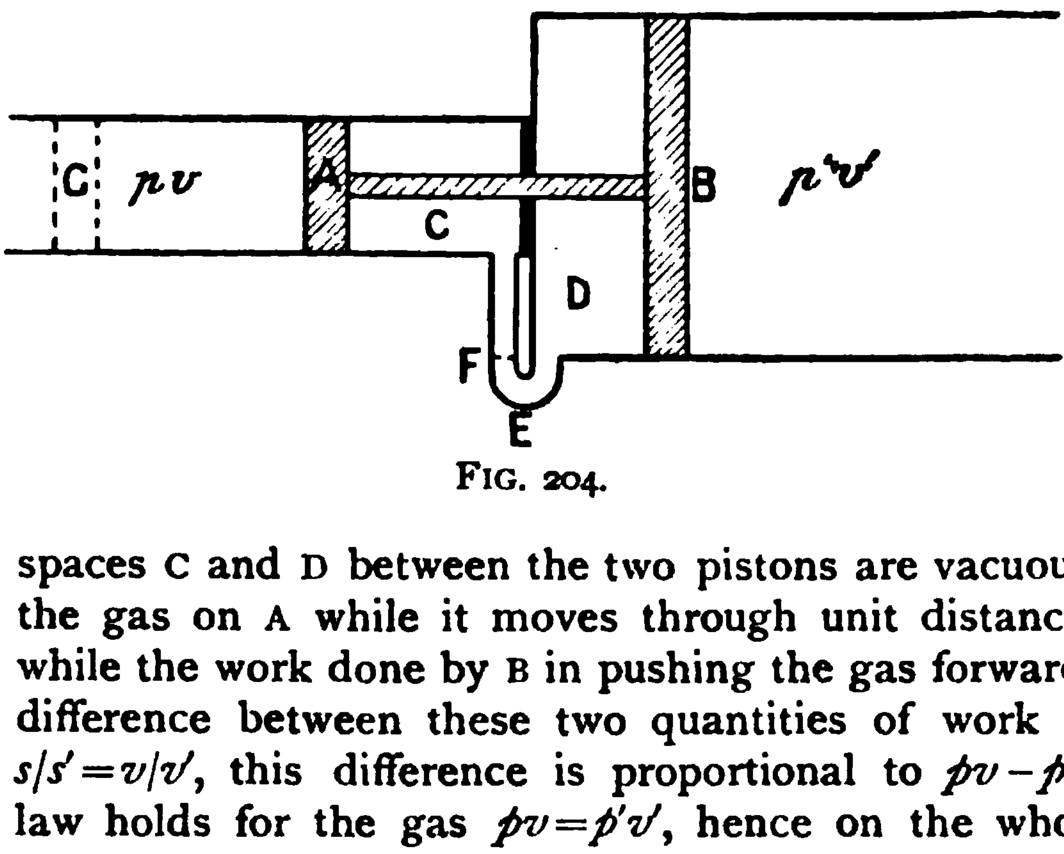


FIG. 204.

the left of A constant, which process might be performed by a second piston, G, working in the cylinder and driven forward by hand. Hence we have allowed the gas to expand without doing external work, and any change of temperature it experiences must be due to the performance of internal work.

The same process would go on if the pistons were not present, for throughout the change we have supposed the pressure on the two sides of each piston to remain the same, so that if gas is allowed to escape through a fine opening, any change in temperature produced will be due to internal action between the molecules. The temperature must not be taken immediately at the opening, for there the gas, as it rushes out, possesses considerable kinetic energy, and it is only after this kinetic energy has been lost by the friction of the gas against itself and against the walls, &c., and the heat energy originally used up in setting the gas in motion is returned to the gas in the form of heat, that no external work has been done on the expanding gas.

In their experiments, Joule and Kelvin allowed a steady stream of gas to pass through a long copper spiral immersed in a water bath kept at a uniform temperature. The gas then escaped through a porous plug made of cotton-wool, which acted the part of the fine hole F, and also prevented the gas from leaving with any appreciable kinetic energy, since the gas rapidly loses its velocity as it passes through the interstices of the wool. The temperature of the gas before and after its passage through the plug was indicated by two delicate thermometers.

In the following table are given some of the results obtained for a difference in pressure between the two sides of the plug of one atmosphere :—

	Temperature before passing through the Plug.	Alteration in Temperature.
	Deg. C.	Deg. C.
Carbon dioxide . .	12.8	− 1.207
”	19.1	− 1.144
”	91.5	− 0.69
Nitrogen . . .	7.2	− 0.305
”	91.7	− 0.187
Oxygen . . .	8.7	− 0.317
”	93.0	− 0.165
Hydrogen . . .	6.8	+ 0.089
”	90.2	+ 0.046
Air	17.1	− 0.255
”	91.6	− 0.203

It will be observed that, except in the case of hydrogen, there is a cooling, indicating that work has to be done in separating the molecules.

The heating obtained in the case of hydrogen may be due to the molecules at ordinary temperatures repelling one another, so that these intermolecular forces do work when the gas expands. It must be remarked, however, that we have supposed that the gas obeys Boyle's law, and that the known deviation from Boyle's law would, in the case of hydrogen, give a heating effect. As the temperature increases the cooling in the case of CO_2, N, O, and air decreases, as also does the heating in the case of H. These results agree with those given in § 130 as to the effects of temperature on the departure of these gases from Boyle's law.

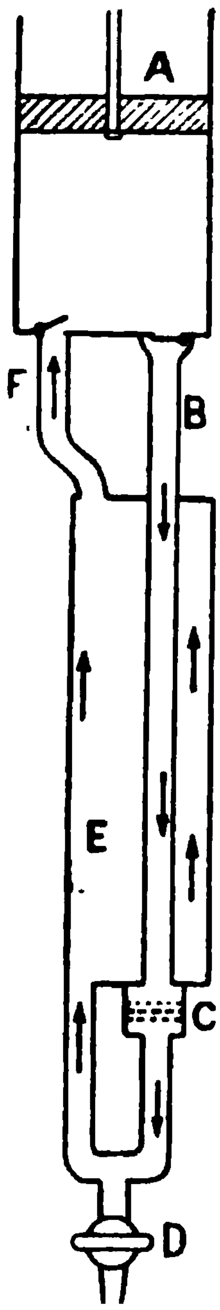


FIG. 205.

The amount of the cooling is proportional to the difference in pressure on the two sides of the plug and to the inverse square of the absolute temperature. Although in Joule and Kelvin's experiment the cooling obtained was so very small, yet by increasing the pressure difference and reducing the temperature, Linde has actually been able to liquefy air by a machine in which this cooling due to the intermolecular forces is used.

A powerful pump A (Fig. 205) draws air out of the tube F and pumps it at a pressure of about fifty atmospheres into the tube B, from which it escapes back into F through a porous plug C. The air becomes cooled in its passage through the plug, and as it passes up through E it cools the descending air in the tube B, which becomes yet colder when passing through the plug C. The gas is pumped round and round in this way, its temperature as it passes down the tube B being always reduced to that of the previous portion when cooled by expansion and so on, till finally the critical temperature is passed and the gas condenses.

255. Relation between Internal and External Work during Change of State.—When a body changes its state, and in doing so changes its volume, the latent heat involved is partly used in doing internal work and partly in external work. Thus in the case of the fusion of ice, when 1 gram of ice at 0° is converted into water at 0° under atmospheric pressure, a contraction of .0907 c.c. takes place, and hence work is done on the body by the atmospheric pressure, and the heat equivalent of this work helps the change of state. The heat equivalent, H, of this work done by the atmospheric pressure is $\frac{Work}{J}=\frac{pv'}{J}$, where p is the atmospheric pressure, and v' the change in volume of unit mass in changing to water. Hence

$$H=\frac{1013300\times .0907}{4.189\times 10^7}=.0022 \text{ calories.}$$

Thus the heat required to perform the internal work necessary to convert 1 gram of ice into water is $L+.0022$ calories, where L is the latent heat.

In the case of steam, the change of volume of 1 gram of water at 100° to steam at 100°, under a pressure of 1 atmosphere, is 1649 c.c. Hence the external work which has to be done by the steam is 1013300×1649 ergs, and the thermal equivalent of this is $\frac{1013300 \times 1649}{4.189 \times 10^7} = 39.9$ calories. Thus the heat spent in internal work is $536.2 - 39.9 = 496.3$ calories.

256*. Theoretical Value of the Difference of the Specific Heats of a Gas.—It has been shown in § 197 that if p and v are the pressure and volume of a mass, m, of a perfect gas at a temperature T, measured from absolute zero, and p_0, v_0 are the pressure and volume at 0° C. or 273° on the absolute scale, we have

$$pv = \frac{p_0 v_0}{273} T.$$

Now if ρ is the density of the gas at a pressure p_0 and at a temperature of 0° C., we have

$$m = v_0 \rho,$$

while p_0 is equal to one standard atmosphere, or, in *c.g.s.* units, 1013260 dynes per square centimetre. Substituting these values,

$$pv = \frac{1013260}{273} \cdot \frac{m}{\rho} T.$$

Or, if we deal only with unit mass of the gas, so that m is unity and v is the volume of unit mass,

$$pv = \frac{3711.6}{\rho} . T.$$

Now as long as we are dealing with any given gas, ρ, the density under standard conditions, is a constant, so that we may write this equation—

$$pv = RT,$$

where R is a constant for any one gas, and is equal to $3711.6/\rho$.

We have seen, in § 253, that

$$J(C_p - C_v) = \frac{p}{\rho T_1},$$

where C_p and C_v are the specific heats of a gas at constant pressure and volume respectively, measured in thermal units (calories per gram), while ρ is the density of the gas at a pressure of p dynes per square centimetre and a temperature T_1 on the absolute scale.

Hence if we take the gas under standard conditions, so that p is equal to 1013260 dynes per square centimetre, and T_1 is 273°, we have

$$J(C_p - C_v) = \frac{1013260}{273\rho}$$
$$= 3711.6/\rho.$$

Thus $J\rho(C_p - C_v)$ is a constant for all perfect gases.

Now
$$R = 3711.6/\rho.$$

Hence
$$J(C_p - C_v) = R.$$

If the specific heats are not measured in thermal units, but in mechanical units, namely ergs, then

$$C_p - C_v = R.$$

Thus the constant R, which appears in the equation

$$pv = RT,$$

is numerically equal to the difference in the specific heats. It must be remembered that the expression

$$J(C_p - C_v) = p/\rho T_1$$

was only obtained on the supposition that there is no force exerted between the molecules of the gas, that is, that $f(v/T)$ is zero. Hence the relation

$$C'_p - C'_v = R,$$

can only be exact for a perfect gas. The following table gives the value of R and of $C'_p - C'_v$ for some gases, and shows to what extent agreement can be expected in the case of an actual gas.

	R.	$C'_p - C'_v$.
Hydrogen	41.3×10^6	41.4×10^6
Air	2.88×10^6	2.87×10^6
Nitrogen	2.97×10^6	2.96×10^6
Carbon dioxide . . .	1.94×10^6	1.88×10^6

It will be noticed that there is a marked difference in the case of carbon dioxide, a gas for which the deviations from Boyle's and Charles's laws is considerable. Also the difference in the case of hydrogen is in the opposite sense to that in the case of the other gases, a result which also agrees with the anomalous behaviour of this gas as regards Boyle's and Charles's laws.

257*. Changes in the Kinetic Energy of the Molecules of a Gas when Heated.—We have seen, in § 142, that if p is the pressure to which a gas is subjected, ρ its density at this pressure and at a tem-

perature T_1, then the mean square of the velocity of translation of its molecules at this temperature is given by

$$p/\rho=\tfrac{1}{3}\overline{V^2}.$$

If v is the volume occupied by 1 gram of the gas at a pressure p and a temperature T_2, we have $v=1/\rho$. Hence

$$pv=\tfrac{1}{3}\overline{V^2}.$$

Now if all the molecules (total mass 1 gram) were moving with the velocity $\overline{V}$, the total kinetic energy K, due to the motion of translation, would be given by

$$K=\tfrac{1}{2}\overline{V^2}.$$

Thus $$pv=\tfrac{2}{3}K.$$

Now suppose that, keeping the pressure constant, the temperature is raised 1°, the volume changing to $v+v^1$, then

$$p(v+v^1)=\tfrac{2}{3}K^1,$$

where K^1 is the total kinetic energy of translation at the temperature $T+1$.

Thus we have $$pv^1=\tfrac{2}{3}(K^1-K).$$

But pv^1 is the external work done by the gas during the expansion. Therefore: the external work done by a gas during expansion is equal to two-thirds of the increase in the kinetic energy due to the motions of translation of the particles.

If a gas is heated from a temperature T_1 to a temperature T_2 at constant volume, and p_1 and p_2 are the original and final pressures, we have, since the density of the gas is the same at both temperatures, and

$$p=\rho\overline{V^2}/3,$$

$$p_1/p_2=\overline{V_1^2}/\overline{V_2^2}.$$

But by Charles's law $$p_1/p_2=T_1/T_2.$$

Hence $$T_1/T_2=\overline{V_1^2}/\overline{V_2^2}$$
$$=\tfrac{1}{2}m\overline{V_1^2}/\tfrac{1}{2}m\overline{V_2^2},$$

where m is the mass of a molecule. Thus the mean kinetic energy of translation of the molecules of a gas is proportional to the temperature.

Thus when the gas is heated at constant volume to $T+1$, the increase in the mean velocity of the molecules will be the same as it was when the gas was heated at constant pressure, the rise in temperature being the same. The increase in kinetic energy of translation will therefore be the same as before.

Now when a gas is heated at constant pressure, omitting the attrac-

tions which the molecules exert one on another, which Joule and Kelvin's experiments (§ 254) have shown to be very small, the energy supplied may be used in three ways—(1) It is employed in doing external work against the external pressure during the expansion; (2) it is employed in increasing the kinetic energy of translation of the molecules; (3) it is employed in increasing the kinetic energy due to the rotation of the molecules as a whole, or to vibrations within the molecules themselves.

Let us first suppose that we have a gas in which all the energy is used up in the first two of these ways.

Then

$$\frac{C_p}{C_v}=\frac{\textit{External work}+\textit{Increment of } K}{\textit{Increment of } K},$$

or, since we have shown that the external work is equal to two-thirds the increment of the translational kinetic energy K,

$$\frac{C_p}{C_v}=\frac{\frac{2}{3}(K^1-K)+(K^1-K)}{K^1-K}$$

$$=\tfrac{5}{3}=1.\dot{6}.$$

Thus, for a gas in which none of the energy is expended in setting up molecular rotation or molecular vibrations, the ratio of the specific heats ought to be 1.667.

Now, if the molecule of a gas consisted of a hard spherical atom, we should expect that there would be no molecular rotation or vibration set up by collisions. If, however, the molecule consists of one or more atoms, which are connected together in some way, we should expect that the collisions would set up vibrations of these atoms within the molecule. Hence when a gas, as is the case for mercury vapour, gives a value for the ratio of the specific heats of 1.667 or thereabouts, we conclude that the molecule of such a gas consists of a single atom.

In a gas in which some of the energy is employed in increasing the molecular rotations and vibrations, the external work is less than two-thirds of the increase of kinetic energy (due to translation, vibration, and rotation).

Thus the fraction

$$\frac{\textit{External work}+\textit{Increase in kinetic energy}}{\textit{Increase in kinetic energy}}$$

is *less* than 1.667.

If, then, the value obtained for the ratio of the specific heats is less than 1.667, we may conclude that the molecule is capable of rotation and vibration, and is therefore complex, and the lower the value of this ratio, the greater the complexity.

In the following table the values of the ratio of the specific heats for some gases and vapour, are given, and it will be noticed that the cases

where the ratio is small are just those in which, from chemical considerations, we should expect a complex molecule.

RATIO OF THE SPECIFIC HEATS.

	Chemical Formula.	Ratio of the Specific Heats.
Mercury	Hg	1.666
Argon	Ar	1.63
Carbon monoxide	Co	1.403
Hydrochloric acid	HCl	1.398
Air	mixture of O_2 and N_2	1.405
Oxygen	O_2	1.41
Hydrogen	H_2	1.41
Nitrogen	N_2	1.41
Chlorine	Cl_2	1.336
Carbon dioxide	CO_2	1.311
Water	H_2O	1.33
Chloroform	$CHCl_3$	1.11
Alcohol	C_2H_6O	1.13
Ether	$C_4H_{10}O$	1.03

In the case of ether, if E is the external work done during an increase of temperature of one degree, when the pressure is kept constant, and k is the *increase* in the total kinetic energy, we have

$$\frac{E+k}{k} = 1.03.$$

Therefore

$$E = .03\,k,$$

so that in this case the external work is only three hundredths of the energy used up in increasing the molecular motions.

We may, if we like, go a step further, and see what proportion the increase in kinetic energy, due to the motion of translation (k_t), bears to the increase in the kinetic energy of vibration and rotation (k_v). For

$$\frac{E + k_t + k_v}{k_t + k_v} = 1.03.$$

Now we have already shown that

$$E = \tfrac{2}{3}k_t.$$

Hence

$$\frac{\tfrac{2}{3}k_t + k_t + k_v}{k_t + k_v} = 1.03,$$

or,

$$k_t = .047\,k_v,$$

so that the energy used in increasing the motion of translation is 4.7 per cent. of the energy used in increasing the motions of vibration and rotation.

258. Adiabatic Curves.—We have in § 231 considered the curves which show the relation between the pressure and volume of a substance when the temperature is kept constant (isothermals), and we have now to consider another set of curves which show the relation between the pressure and volume of a substance when these quantities change, but no *heat* is allowed to enter or leave the substance. These curves are called *adiabatics*.

It follows from what has been said in § 252 that when a body expands and does external work, then, if no heat is supplied to the body, its temperature must fall, while if under the same conditions the body is compressed, the temperature will rise. If AB (Fig. 206) represents a portion of the isothermal through a point P, then if, starting with the substance in the conditions indicated by P, we compress the substance adiabatically, we do work on it, and therefore its temperature will rise, and for a given pressure the volume will be greater than it would be if we had kept the temperature constant, *i.e.* travelled along the isothermal PA. Hence the adiabatic CD through P is more steep than the isothermal through the same point.

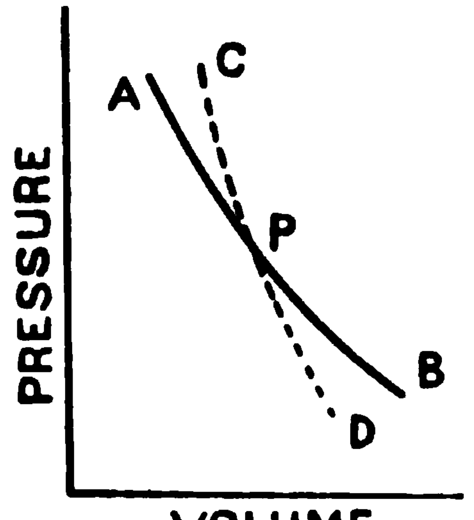


FIG. 206.

The equation to an adiabatic curve for a perfect gas is $pv^k =$ constant, where k is the ratio of the specific heat at constant pressure to the specific heat at constant volume. The proof of this equation would, however, lead us too far into the subject, and we must content ourselves with simply stating the fact.

259*. Direct Determination of the Ratio of the Specific Heats for a Gas.—The most direct way of measuring the ratio of the specific heat at constant pressure to that at constant volume is that employed by Clément and Desormes.

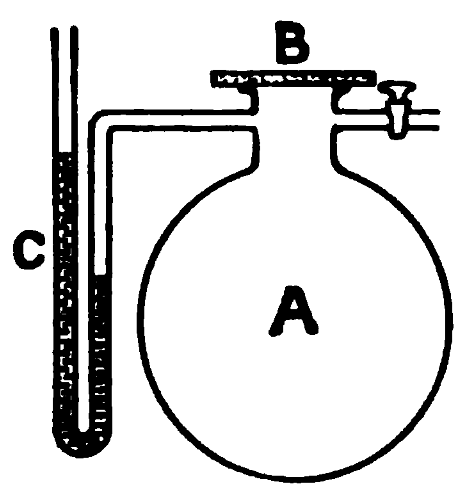


FIG. 207.

The gas to be experimented on is contained in a large glass balloon A (Fig. 207), which has a wide mouth that can be closed by a plate of ground glass B. A manometer C serves to measure the pressure of the gas. The opening B being closed, a little of the gas is pumped in, so as to make the pressure p_1 a little greater than the atmospheric pressure. After the heating caused by the compression has been dissipated by conduction, &c., the plate B is removed just long enough to allow the pressure inside to fall to the atmospheric pressure p_0 by some of the gas escaping. The gas inside the globe expands, and the expansion is so

rapid that no appreciable quantity of heat has time to pass from the walls to the gas, so that the expansion is adiabatic, and the temperature of the air is lowered. After the closing of the opening, however, the air becomes gradually heated to its old temperature by heat derived from the walls of the vessel, and hence the pressure p_2 rises above the atmospheric pressure. If v_1 is the volume of *unit mass* of the gas when compressed under the pressure p_1, and v_2 is the volume of unit mass after expansion, we have during the adiabatic expansion that

$$p_1 v_1^k = p_0 v_2^k.$$

Hence

$$\frac{p_0}{p_1} = \left(\frac{v_1}{v_2}\right)^k.$$

Also, since the first temperature of the gas, when the pressure and volume were $p_1 v_1$, and the last temperature, when the pressure and volume are $p_2 v_2$, are the same, we have by Boyle's law—

$$p_1 v_1 = p_2 v_2.$$

$$\therefore \quad \frac{v_1}{v_2} = \frac{p_2}{p_1}.$$

Substituting we get

$$\frac{p_0}{p_1} = \left(\frac{p_2}{p_1}\right)^k,$$

or, taking logarithms of both sides,

$$k = \frac{\log p_0 - \log p_1}{\log p_2 - \log p_1}.$$

Hence, from the observed pressure, p_1, p_0, and p_2, the value of k can be calculated.

The results obtained by this method are not very trustworthy, since it is difficult to secure perfectly adiabatic expansion, that is, without appreciable passage of heat from the walls to the gas during the time it is expanding, and at the same time prevent the outflow of the gas being oscillatory, so that the pressure is alternately less and greater than p_0, and only becomes steady after a few oscillations.

Another method of determining the value of k will be described when we are considering the velocity of sound in gases.

260. Watt's Indicator Diagram.—In § 251 we have seen that, when a gas expands at constant pressure against a pressure p, the external work done is equal to pv', where v' is the change in volume. Suppose we start with a given quantity of a gas having a volume v_1 under a pressure p_1, as indicated by the point A (Fig. 208) on the diagram of pressures and volumes, then if the gas expands at constant pressure to the volume v_2, as indicated by the point B, the change that takes place is represented by the horizontal line AB. The external work done is $p_1(v_2 - v_1)$, and since p_1 is represented by the line $\overline{\text{A}v_1}$ or $\overline{\text{B}v_2}$, we see

that $p_1(v_2 - v_1)$ is numerically equal to the area of the rectangular figure $\overline{\text{AB}v_2v_1}$. Hence the external work done by the gas is represented by the area of this figure. If now the pressure is decreased to p_2, the volume being kept constant (of course, the temperature will have to be lowered), the change is indicated by the line BC, and since the volume does not change, no external work is done. Next, keeping the pressure constant, let the gas be cooled till its volume is again v_1, the change being represented by the line DC. During this process work must be done *on* the gas, and the amount of this work is represented by the area of the rectangle $\overline{\text{DC}v_2v_1}$. Finally, keeping the volume constant, heat the gas till the pressure rises to p_1, the line DA representing the new change, which is unaccompanied by any external work. The gas is now in exactly the same state as it was at the start, and we have taken it through what is called a cycle of operations. During this cycle the gas has performed an amount of external work represented by the figure $\overline{\text{AB}v_2v_1}$, and had an amount of work represented by $\overline{\text{DC}v_2v_1}$ done on it. Hence the total amount of external work done by the gas during the cycle is represented by the rectangle $\overline{\text{ABCD}}$ enclosed by the path which indicates the different conditions to which the gas has been subjected during the cycle.

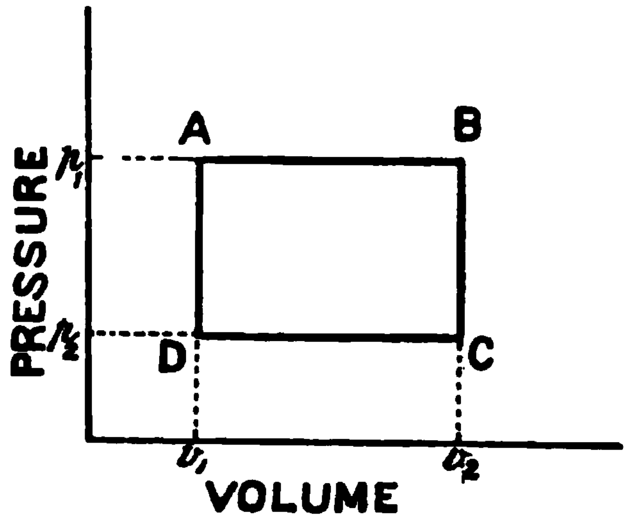


FIG. 208.

If the cycle had been traversed in the opposite sense, namely, ADCBA, it can easily be shown that on the whole an amount of work represented by the rectangle $\overline{\text{ABCD}}$ would have been done *on* the gas. Hence, if such a cycle is traversed in the clockwise direction work is done by the gas, and if in the anticlockwise direction work is done on the gas. In the above example we have supposed that the pressure and volume changed one at a time. If, now, we suppose them to change simultaneously, and the gas to change from the conditions indicated by the point A (Fig. 209) to those indicated by the point B along the curve $\overline{\text{AB}}$, we may suppose that the curve is replaced by a stepped curve such as that shown, along which the pressure and volume only change one at a time. When the gas is going from A to C, the volume being constant, no work is done. When the part CD is being traversed, the work

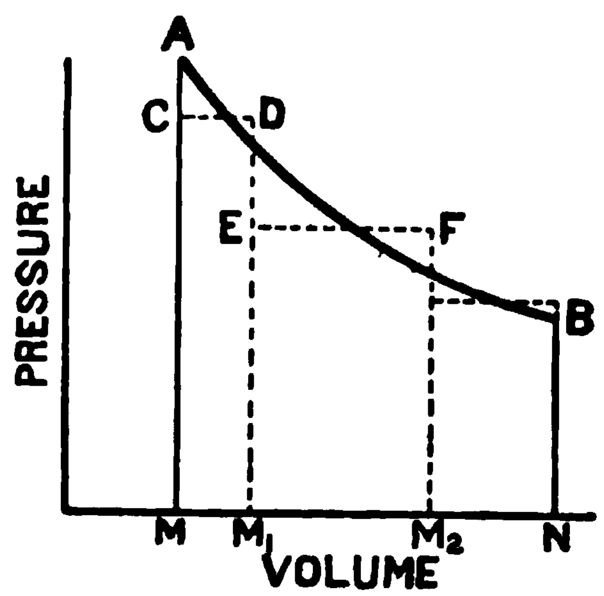


FIG. 209.

done is represented by the rectangle $\overline{CDM_1M}$, and so on. Now if we imagine the number of steps taken as infinitely increased, the stepped curve will nowhere appreciably differ from the curve $\overline{AB}$,[1] and the work done will be represented by the area ABNM included between the curve, the axis of volumes, and the two ordinates drawn through the extreme points A and B. Hence it follows that in the case of a closed cycle, the work done by the body, if the path is traversed in the clockwise direction, or the work done on the body, if the path is traversed in the anti-clockwise direction, is equal to the area enclosed by the path which represents the cycle of operations.

261*. Carnot's Cycle.—A particular cycle, which is found to be of great use in the theory of heat, is one due to Carnot. We will suppose

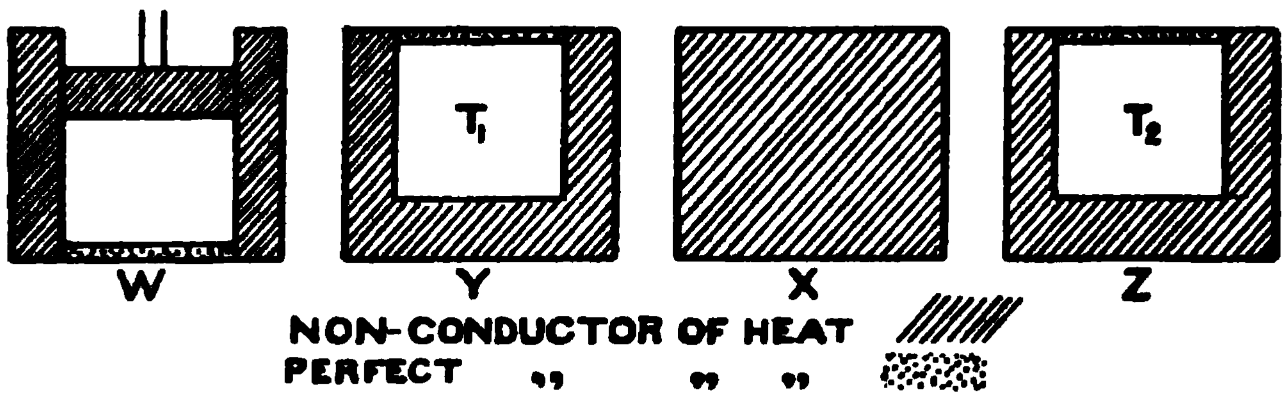


FIG. 210.

that the substance which we are about to cause to go through a Carnot's cycle (the working substance, as it is called) is contained within a cylinder W (Fig. 210), the walls and piston of which are perfect non-conductors for heat, and that the bottom of the cylinder is made of a perfect conductor of heat. Further, we will suppose that we have three stands, one, X, fitted with a perfect non-conducting top, and of the others one, Y,

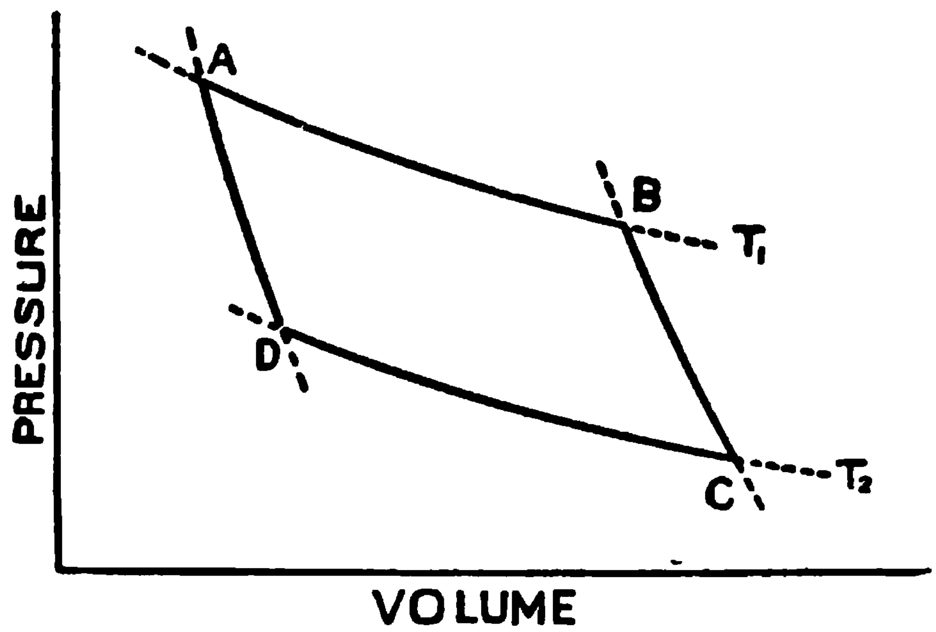


FIG. 211.

kept at a constant temperature T_1, and the other, Z, at a constant temperature T_2, these being each fitted with a perfectly-conducting top.

First place the cylinder on the stand Z, so that the working substance comes to a temperature T_2, the pressure and volume being as indicated by the point D (Fig. 211). Now place the cylinder on the non-conducting

[1] The reasoning is the same as that adopted in § 34.

stand X, and increase the pressure, and so decrease the volume. The change must be adiabatic, since the non-conducting stand prevents the escape of the heat due to the compression. The compression must be stopped when the temperature has risen to T_1. The curve DA, which gives the relation between the pressure and volume during this operation, is an adiabatic (§ 258). Next place the cylinder in the stand Y, and expand the substance by allowing the piston to rise till the volume and pressure are indicated by the point B. During this operation heat will flow into the working substance through the conducting bottom of the cylinder, so that the temperature will be constant throughout the process, and hence the curve $\overline{AB}$ will be a portion of the isothermal for the temperature T_1. Next remove the cylinder from Y, and place it on the non-conducting stand X, and continue allowing the substance to expand till the temperature falls to T_2. This portion of the cycle, since the escape or supply of heat to the working substance is prevented by the non-conducting stand, is adiabatic. Next place the cylinder on Z, and force the piston down till the pressure and volume become the same as at the start. During this part of the operation heat is given out by the working substance to the stand, and $\overline{CD}$ is a portion of the isothermal for T_2. The cycle is now complete, the working substance being in exactly the same condition as at the start. During the cycle an amount of work represented by the area ABCD has been performed, also during the portion of the cycle represented by $\overline{AB}$ heat has been supplied to the working substance at a temperature T_1, and during the portion CD heat has been given out by the working substance at the temperature T_2. If H_1 is the quantity of heat taken in at the temperature T_1, and H_2 the quantity of heat given out at the temperature T_2, then from the first law of thermodynamics it follows, since the initial and final states of the working substance are the same, that the work W done during the cycle must be equivalent to the heat which has been used, so that

$$W = J(H_1 - H_2).$$

Now the maximum quantity of work which can be obtained from H_1 units of heat is JH_1. The ratio of the actual amount of work done in any cycle to the maximum amount of work that could be done, suppose all the heat supplied had been converted into mechanical energy, is called the efficiency of the cycle. Hence the efficiency (n) is given by

$$n = \frac{H_1 - H_2}{H_1}.$$

The peculiarity of Carnot's cycle is that the cycle may be traversed in the reverse direction, a quantity of heat H_2 being taken in by the working substance at a temperature T_2, and a quantity of heat H_1 given out at a temperature T_1, while a quantity of work, represented by the area ABCD,

has to be done on the working substance during the cycle. For this reason Carnot's cycle is called a reversible cycle.

Carnot also showed that by adopting a reversible cycle the efficiency obtained was the greatest possible. For if not, suppose we had two engines working between the same temperatures T_1 and T_2, one, A, working in a reversible cycle, and the other, B, having a greater efficiency than A, and that B works direct, taking in heat from a body at a temperature T_1, and giving out heat to a cold body at a temperature T_2, and that the mechanical work it does is employed in working the reversible engine A backwards, so that it takes in heat at a temperature T_2, and gives it out at a temperature T_1. Let the quantity of heat taken in by B at the higher temperature be H_1, and that given out H'_2, while the heat taken in by A at the lower temperature be H_2, and that given out at the higher temperature be H_1. Then $H'_1 - H'_2$ is the heat converted into work by B, and $H_1 - H_2$ is the work converted into heat by A. Now, if all the work done by B is used in working A, it follows that $H'_1 - H_2 = H_1 - H_2$. The efficiency of A is $\frac{H_1 - H_2}{H_1}$, and that of B is $\frac{H'_1 - H'_2}{H'_1}$, hence by supposition $\frac{H_1 - H_2}{H_1}$ is less than $\frac{H'_1 - H'_2}{H'_1}$. Therefore since $H_1 - H_2$ is equal to $H'_1 - H'_2$, it follows that H_1 must be greater than H'_1, so that H_2 must also be greater than H'_2. In other words, the heat H_2 taken from the cold body by A is greater than the heat H'_2 supplied to it by B, while the heat H'_1 supplied to the hot body by A is greater than the heat H_1 taken from it by B. Thus as the combination continues to work the cold body will be gradually robbed of all its heat, while the heat of the hot body will increase, and this without the expenditure of any external energy, which is entirely contrary to experience; for, unless there is an expenditure of external energy, experience shows that heat always flows from the body at the higher temperature to that at the lower. Hence we are led to the conclusion that an engine working in a Carnot's reversible cycle is the most efficient that it is possible to have working between the two given temperatures.

It also follows that since a substance when it goes through a Carnot's reversible cycle, taking in a quantity of heat H_1 at a temperature T_1, and giving out heat at a temperature T_2, converts the maximum fraction of the heat received into work, that all working substances when used in a Carnot's reversible cycle must have the same efficiency. That is, that the fraction (work done during cycle)$\div H_1$ must be a constant, so that so long as H_1, the quantity of heat taken in, remains the same, the quantity of work, W, done during the cycle must remain the same, and depends only on the temperatures T_1 and T_2.

At first sight the reason for this may not appear quite clear. What we show is that for an engine working in a reversible cycle the efficiency is the highest possible whatever the working substance, for we have

made no assumptions as to the nature of the working substance. Now if the efficiencies of a number of reversible engines with different working substances are *all* the maximum, they must all be equal.

The above result may not appear to agree with common sense, for if the working substance in an engine is ether, say, we might at first sight hope to obtain a higher efficiency than with water. For the vapour pressure of ether at any temperature being higher than that of water at the same temperature, the pressure in a boiler filled with ether would be greater than in a boiler filled with water at the same temperature. Thus the vapour supplied to the ether engine will be at a higher pressure than that supplied by the water engine, and so we might expect that we should get more work by allowing the ether vapour to force back a piston than in the case of the steam. It must, however, be remembered that at the back of the piston we have acting the pressure of the exhaust steam or vapour, and that the condenser for the two engines must be supposed to be at the same temperature, or they would not be working through the same range. Now the vapour pressure of the ether at the temperature of the condenser will be greater than that of the water, and so not only the forward pressure in the ether engine but also the back pressure is greater than in the water engine, and so no advantage is gained.

Hence a Carnot's cycle being a thermal process which is independent of the nature of the substance in which the thermal changes take place, it at once becomes of interest to see whether we cannot utilise this fact in order to define a scale of temperature independent of all properties of any particular kind of matter. The scales which we have used heretofore all depend on the change of some one physical property of some special kind of matter, thus on the increase in volume of mercury or hydrogen, the increase in resistance of a platinum wire, the thermo-electromotive force of a junction of two given metals, &c. A scale of temperature depending on Carnot's cycle and independent of the properties of any particular kind of matter has, however, been devised by Lord Kelvin, and to this scale only can the title "absolute" be given with justice.

If, as before, we imagine a Carnot's cycle in which a quantity of heat H_1 is drawn from a source at a temperature T_1, and an amount of work W is performed, H_2 units of heat being given out to the refrigerator at the temperature T_2, we may according to the first law measure H_2 and H_1 in terms of ergs, in which case $H_1 - H_2 = W$. If now, keeping H_1 and T_1 constant, we adjust the temperature T_2 so that the work done during the cycle is unity, then the two temperatures T_2 and T_1 will be such that if a Carnot's engine working between these temperatures takes H_1 ergs of heat from the source it will perform one erg of work. Next suppose that another cycle is taken, in which the lower temperature T_3 is so adjusted that when H_1 ergs of heat are drawn from the source at a temperature T_1, the work done in the cycle is two ergs. Then, according to Lord Kelvin, the difference of temperature between T_1 and T_3 is to be

called twice the difference of temperature between T_1 and T_2. Proceeding in this way, we could define a series of equal temperature intervals, and thus form a thermometric scale. It will, however, be convenient not to call the interval $T_1 - T_2$, or $T'_1 - T'_2$, as above defined, one degree, since the scale thus constructed would not resemble the scale ordinarily employed. We will therefore suppose that T'_1 is taken as the temperature of boiling water, and we will postulate that when H_1 units of heat are taken, by an engine working in a simple reversible cycle, from a source at the temperature of boiling water, and the refrigerator is at the temperature of melting ice, a hundred times the work will be done that would be done supposing the temperature of the refrigerator were one degree, on this new absolute scale, below the temperature of boiling water, and so on.

Let the lines T_0T_0 and $T_{10}T_{10}$ (Fig. 212) be the isothermals for the temperatures of melting ice and boiling water respectively, and let AB be an adiabatic cutting these isothermals at E and G. Suppose that if we go along the isothermal T_{10} from E to F an amount of heat H_{10} (measured in ergs) has to be supplied to the working substance to keep its temperature constant, and that through F we draw a second adiabatic CD cutting the isothermal T_0 at H. Then, if a simple reversible engine performs the cycle EFGH, it will take in H_{10} units of heat at a temperature T_{10} and give out H_0 units of heat at a temperature T_0, while the work W_{10} done during the cycle will be represented by the area of the figure EFHG. Now draw

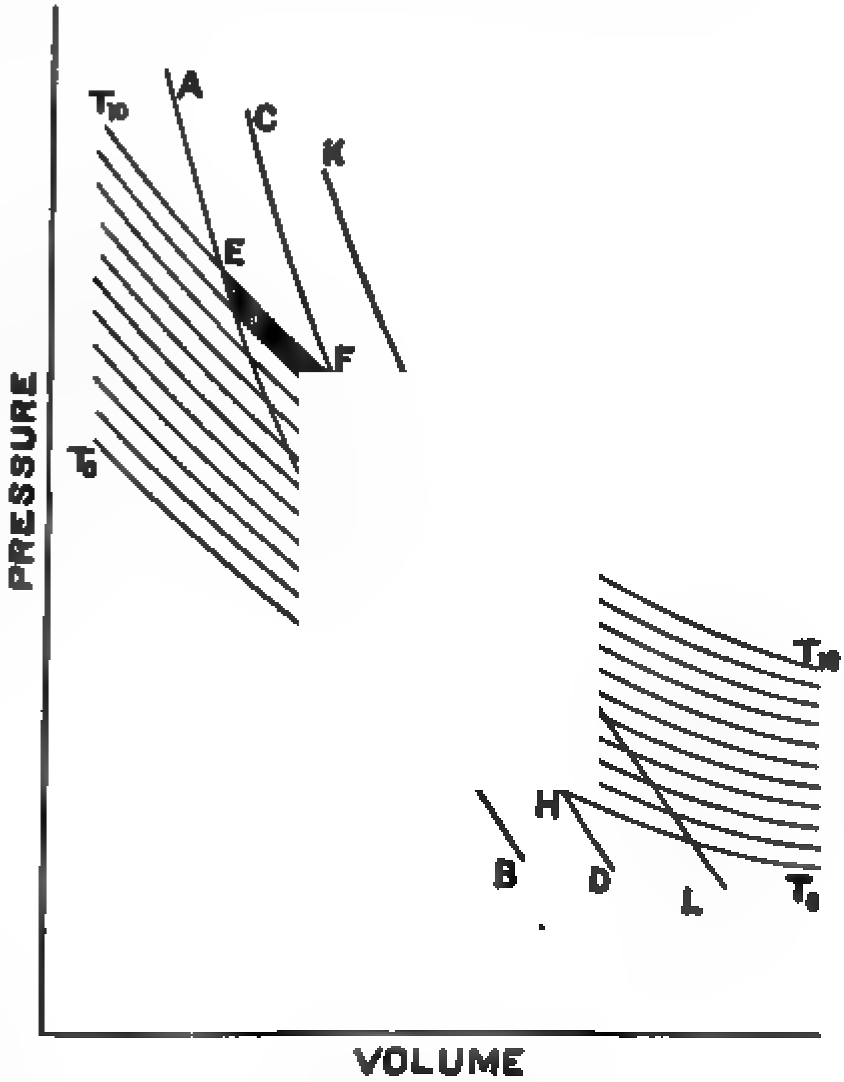


Fig. 212.

nine isothermals between T_0 and T_{10}, so spaced that the area intercepted between any adjacent two and the two adiabatics is one-tenth of the area EFGH. Thus the area shown shaded is to be one-tenth of EFGH. Then the temperatures corresponding to these isothermals, if we call the temperature of melting ice 273° and that of boiling water 373°, are 283°, 293°, 303°, 313°, 323°, 333°, 343°, 353°, 363° on Lord Kelvin's absolute scale.

By the doctrine of the conservation of energy, the maximum amount of work we can possibly get from a quantity of heat H_1 is JH_1, if the quantity H_1 is expressed in calories, or simply H_1, if this quantity is expressed in ergs. Keeping the temperature T_1 of the source constant, the amount of work W obtained during a cycle will increase as the temperature of the refrigerator is lowered, until the temperature of the refrigerator becomes such that no heat is given to it during the compression portion of the cycle, the whole of the heat taken in being converted into work, so that $H_1 = W$, or the efficiency of the cycle becomes unity. We cannot imagine a refrigerator at a lower temperature than this, and hence may take it as the zero on this new absolute scale. It is found that the zero thus defined coincides with the absolute zero as given by a perfect gas, and that the new absolute scale agrees very nearly with that of a gas thermometer containing a perfect gas. So that the use of the thermometric scale derived from the expansion of a perfect gas is justified.

If, in Fig. 212, any other adiabatic KL is drawn, then this, together with either of the others, will cut off equal areas between consecutive isothermals. Thus the area intercepted by any two adiabatics and any two isothermals T_1 and T_2, say, will be proportional to the difference of temperature $T_1 - T_2$, for each degree of this difference will correspond to an equal small area k, such as the one shown shaded. Thus we shall have

$$W = k(T_1 - T_2),$$

where k is a constant depending on the two adiabatics taken. Since $W = H_1 - H_2$, this gives

$$H_1 - H_2 = k(T_1 - T_2).$$

Now if we make T_2 the absolute zero, there will be T_1 small areas each equal to k included in the cycle, and H_2 will be zero, so that in this case

$$W = H_1 = kT_1.$$

Now the efficiency of a reversible cycle is given by

$$n = \frac{H_1 - H_2}{H_1}.$$

Hence, substituting for $H_1 - H_2$ and H_1,

$$n = \frac{H_1 - H_2}{H_1} = \frac{k(T_1 - T_2)}{kT_1} = \frac{T_1 - T_2}{T_1}.$$

This result will be found useful when we are considering some actual cases of reversible cycles.

The above equation may be written

$$1 - \frac{H_2}{H_1} = 1 - \frac{T_2}{T_1},$$

so that

$$\frac{H_1}{H_2} = \frac{T_1}{T_2},$$

or the ratio of the heat taken from the source by a reversible engine to the heat given up to the refrigerator is the same as the ratio of the temperature, on the absolute scale, of the source to that of the refrigerator.

262*. The Second Law of Thermo-Dynamics.—When considering the efficiency of a simple reversible engine, we said that the transfer of the heat of the condenser to the source was contrary to experience. The denial of the possibility of any such action forms what is called the second law of thermo-dynamics,[1] and has been put into a concise form by Clausius, who expresses it as follows: *It is impossible for a self-acting machine, unaided by any external agency, to convey heat from one body to another at a higher temperature.*

Lord Kelvin has enunciated the second law in a slightly different form, namely: *It is impossible, by means of inanimate material agency, to derive mechanical effect from any portion of matter by cooling it below the temperature of the coldest of surrounding bodies.*

It must be carefully borne in mind that these laws refer only to the work performed during a *cycle* of operations, in which the initial and final states of the working substance are exactly the same. Thus when a gas is allowed to expand against external pressure, it does work and becomes cooled, so that in this way it may do work although in the operation it becomes cooled below the temperature of surrounding objects. The final state of the gas is not, however, the same as the initial state, and if we attempt to bring the gas back into the initial state we shall find that the law holds.

We may also put the law in slightly different words, viz. that heat of itself never passes from one body to another at a higher temperature; and if by any means we cause heat to be transferred from a body to another at a higher temperature, we must in the process supply the system with energy from some outside source. Thus, when a reversible engine is worked backwards, heat is taken from the refrigerator and supplied to the source. During this operation, however, external energy has to be supplied to the engine, so that it is not working "by itself."

263*. Calculation of the Effect of an Increase of Pressure on the Melting-Point of Ice.—The second law of thermo-dynamics will allow us to calculate the effect of pressure on the melting-point of ice.

[1] The deductions made in the last section are also generally referred to as forming part of the second law.

Suppose we have a gram of water at 0° C. and at a pressure of one atmosphere, the conditions being represented by the point A, Fig. 213. Now allow the water to freeze. During this process the temperature and pressure will remain constant, so that the horizontal line AB will represent the change, which is an isothermal one. During this change 80 calories of heat, or, if we use mechanical units, $80 \times J$ ergs will be given out, and an amount of work pv', where v' is the change of volume, will be done. One atmosphere being 1013300 dynes per square centimetre, and v' being in the case before us 0.0907, the work done is

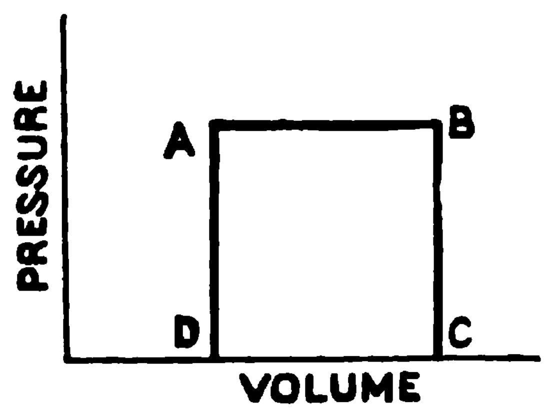


FIG. 213.

$$1013300 \times 0.0907 \text{ ergs.}$$

Now, without allowing heat to enter or leave the ice, reduce the pressure to zero. Since the change in volume of ice (or water) with a difference of one atmosphere is quite inappreciable, the line BC showing this change is vertical, and no work is done, supposing that during the change the ice does not melt, that is, if the melting-point of ice does not vary with the temperature. Next, supply heat to the ice so that it melts, and we now pass along CD. During this process heat is absorbed, but since the pressure is zero, although the volume decreases, no external work is done on the working substance. Finally, raise the pressure to one atmosphere along the adiabatic DA. We have now gone through a reversible cycle of operations in which an amount of work represented by the area ABCD has been done. On the supposition, however, that the melting-point of ice is the same at a pressure of one atmosphere as in a vacuum, the temperature at which the heat was taken in is the same as that at which the heat was given out to the refrigerator. This is, however, contrary to the second law of thermo-dynamics, and hence we conclude that our supposition that the melting-point of ice is unaltered by change of pressure must be wrong. Since external work is done during the cycle, the temperature when the heat was being taken in by the working substance, that is, while the ice was melting at the low pressure along CD, must have been *higher* than the temperature when the heat was being given out, that is, when the water was freezing at the higher pressure. In short, decreasing the pressure has raised the melting-point.

We may proceed to calculate what would be the rise in the melting-point produced by an increase of one atmosphere. Let t be the difference in the temperature of melting ice produced by a change in pressure of one atmosphere. The temperature of the refrigerator is 0° C., or 273° on the absolute scale, and that of the source $273 + t$. The heat absorbed is 80 calories or 3.352×10^9 ergs, and the work done is $0.0907p$ ergs,

where p is one atmosphere expressed in dynes per square centimetre. Now (§ 261)

$$\frac{W}{H}=\frac{T_1-T_2}{T_1}.$$

Therefore
$$\frac{0.0907p}{3.352\times10^9}=\frac{t}{273+t}.$$

Now t is very small compared to 273, so that we shall not produce any appreciable error in omitting the term t in the denominator of the right-hand member. Thus

$$0.0907\times273\times p=3.352\times10^9 t$$
$$p=1.35\times10^8 t.$$

If p is one atmosphere, or 1013260 dynes per square centimetre,

$$t=\frac{1013260}{1.35\times10^8}$$
$$=0^\circ.0075.$$

This number agrees with the results of experiment.

264*. Irreversible Cycles.—The cycles which we have up to now considered have all been reversible, that is, if they are worked backwards, so that all the various operations are performed in the reverse order and sense, the physical and mechanical changes are also reversed. There are, however, many cycles of operations in which, for various reasons, the operations cannot be reversed, or, if they are, the mechanical changes are not reversed. Thus if during any cycle any of the energy is employed in producing motion against friction, such a cycle cannot be reversible, for, as we have seen in § 110, although we reverse the direction of motion, the conversion of mechanical energy into heat due to friction always takes place. Thus, when working direct, the engine working in such a cycle may do a certain quantity of mechanical work owing to the expenditure of a certain quantity of heat-energy ; yet if we reverse the engine and do work on it to the same extent as it did before, since some of this energy is employed in doing work against friction, we shall not completely reverse the thermal processes. Again, if during any part of a cycle there is conduction of heat from one part of the engine to any other, since heat *only* flows by conduction from bodies at higher temperatures to those at lower, on reversing the engine the heat that passed by conduction from a higher to a lower temperature is not made to pass in the reverse direction. It was to avoid the conduction of heat that, in describing Carnot's reversible cycle, we had to suppose that the walls of the cylinder were composed of a perfect non-conductor of heat. Also during a reversible process, when there is passage of heat from one body to another, as, for instance, in the Carnot cycle during the isothermal expansion, when the working substance is taking heat from the

source, it is necessary to suppose that the transference takes place so slowly that the temperature of the working substance never differs by more than an infinitesimal amount from that of the source. If this were not so, when we reversed the cycle, in order to reverse the conditions exactly, we should still require to have the temperature of the working substance higher than that of the source by the same amount as before, and yet have heat flowing from the source to the working substance, *i.e.* from a cold to a hot body.

265. Dimensions of Thermal Quantities.—We have used two distinct units of quantity of heat. One of these, the calorie, depends on the thermal capacity of water, and on the scale of temperature adopted, as well as on the unit of mass. The other, the erg, simply depends on the fundamental units of mass, length, and time, and has the dimensions $[ML^2T^{-2}]$. If Q represents a certain quantity of heat measured in calories, and H the same quantity measured in ergs, then by the first law of thermo-dynamics we have

$$H = JQ,$$

where J is the mechanical equivalent.

Therefore $$[H] = [ML^2T^{-2}] = [JQ].$$

Now in § 251 we have seen that the value of J depends on the scale of temperature adopted, since the value of Q depends on this scale. Hence the dimensions of J depend on the temperature scale. We do not, however, know the dimensions of temperature, as measured on the ordinary gas-thermometer scale, in terms of the fundamental units of length, mass, and time, and so we are reduced to using a symbol $[\theta]$ for the unknown dimensions of temperature. Since the thermal unit depends on the mass of water taken, as well as on the unit of temperature, we have

$$[Q] = [M\theta].$$

Hence $$[J] = [Q^{-1} \, . \, ML^2T^{-2}]$$
$$= [L^2T^{-2}\theta^{-1}].$$

The symbol θ here plays the part of a fourth fundamental unit, and Professor Rücker has proposed to call it a *secondary fundamental unit.* There is no doubt that it is only the limit of our knowledge as to the nature of temperature which prevents our expressing $[\theta]$ in terms of $[L]$, $[M]$, and $[T]$. For instance, we have in § 257 supposed that in the case of a gas the temperature is proportional to the mean kinetic energy of translation of the molecules. Hence we might measure temperatures by the mean kinetic energy of a molecule of a gas when at that temperature, and we should on this scale have

$$[\theta] = [ML^2T^{-2}].$$

As yet, such a method of measuring temperature is not warranted by our knowledge of the molecular conditions of gases, to say nothing of liquids and solids. It is, therefore, better to retain, when dealing with dimensional formulæ involving temperature, the symbol $[\theta]$ for the dimension of the unit of temperature.

Since specific heat is the quantity of heat required to raise unit mass through a temperature of one degree,

$$Q=[Ms\theta],$$

or

$$[s]=[M\theta]/[M\theta]=1.$$

So that specific heat has no dimension, and is therefore a mere number. This is at once evident, if we remember that specific heat may also be defined as the ratio of the heat required to raise a given mass of the substance through a given range of temperature, to the heat required to raise an equal mass of water through the same range.

Latent heat being the quantity of heat required to convert unit mass of the substance from one state to the other,

$$[L]=[Q]/[M]=[\theta].$$

Since the coefficients of expansion are

$$\frac{\textit{Increase of length (or volume)}}{\textit{Original length (or volume)}}\cdot\frac{1}{\textit{Increase in Temperature}},$$

we have

$$[a]=[L]/[L\theta]=[\theta^{-1}]$$
$$[\gamma]=[L^3]/[L^3\theta]=[\theta^{-1}].$$

The quantity of heat, Q, which passes in a time, t, through a slab of area A, thickness d and conductivity k, when the difference of temperature between the opposite faces is θ, is given by

$$Q=kt\theta A/d.$$

Therefore

$$[k]=[QL]/[T\theta L^2]$$
$$=[L^{-1}MT^{-1}].$$

BOOK III

WAVE-MOTION AND SOUND

PART I—WAVE-MOTION

CHAPTER I

WAVE-MOTION AND WATER WAVES

266. Wave-Motion.—We have in Book I. chap. vii. considered the periodic motion of a single particle or rigid body; we have now to consider in some detail the resultant motion when the various particles of a medium are executing periodic motions, but the phase (§ 50) of the motions of the various particles is not the same for all, but are related to one another in certain definite ways.

Suppose we have a number of particles arranged, when at rest, at equal distances along a line AB (Fig. 214), and that these particles all

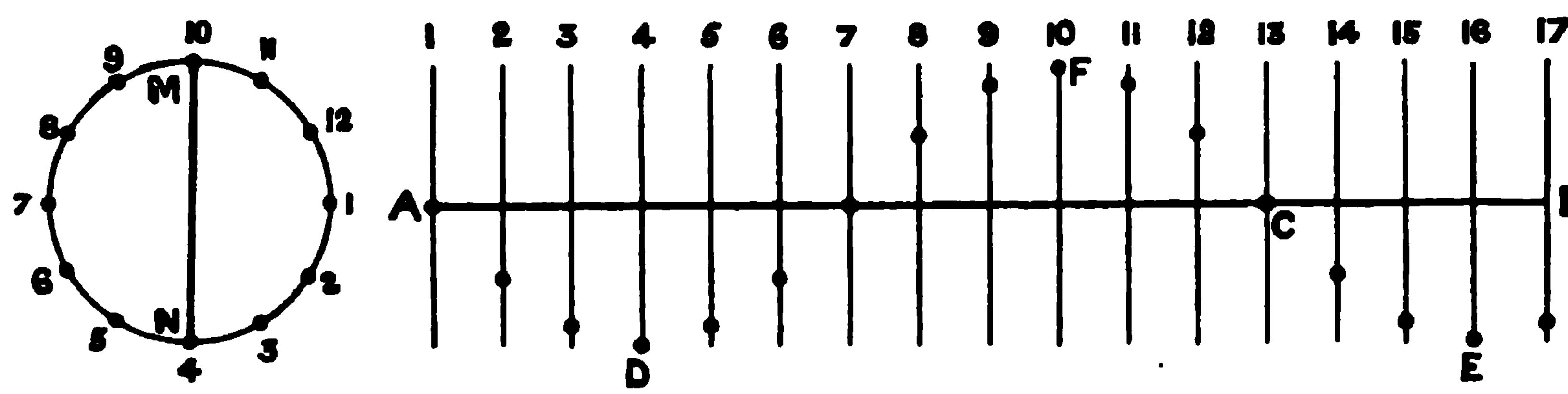


FIG. 214.

execute S.H.M.'s (§ 50) of equal amplitude and period along lines at right angles to AB, but in such a way that the phase of each successive particle, counting from A, differs from that of the preceding particle by a constant amount.

Thus if the constant difference in phase is 30°, when the particle 1 is at its median position, the position of the others will be as shown by the dots in the figure. The displacement of particle 2 at any moment is equal to the displacement of particle 1 at 1/12 of the periodic time (T)

later, since 30° is 1/12 of 360°. Similarly, particle 3 is displaced to the

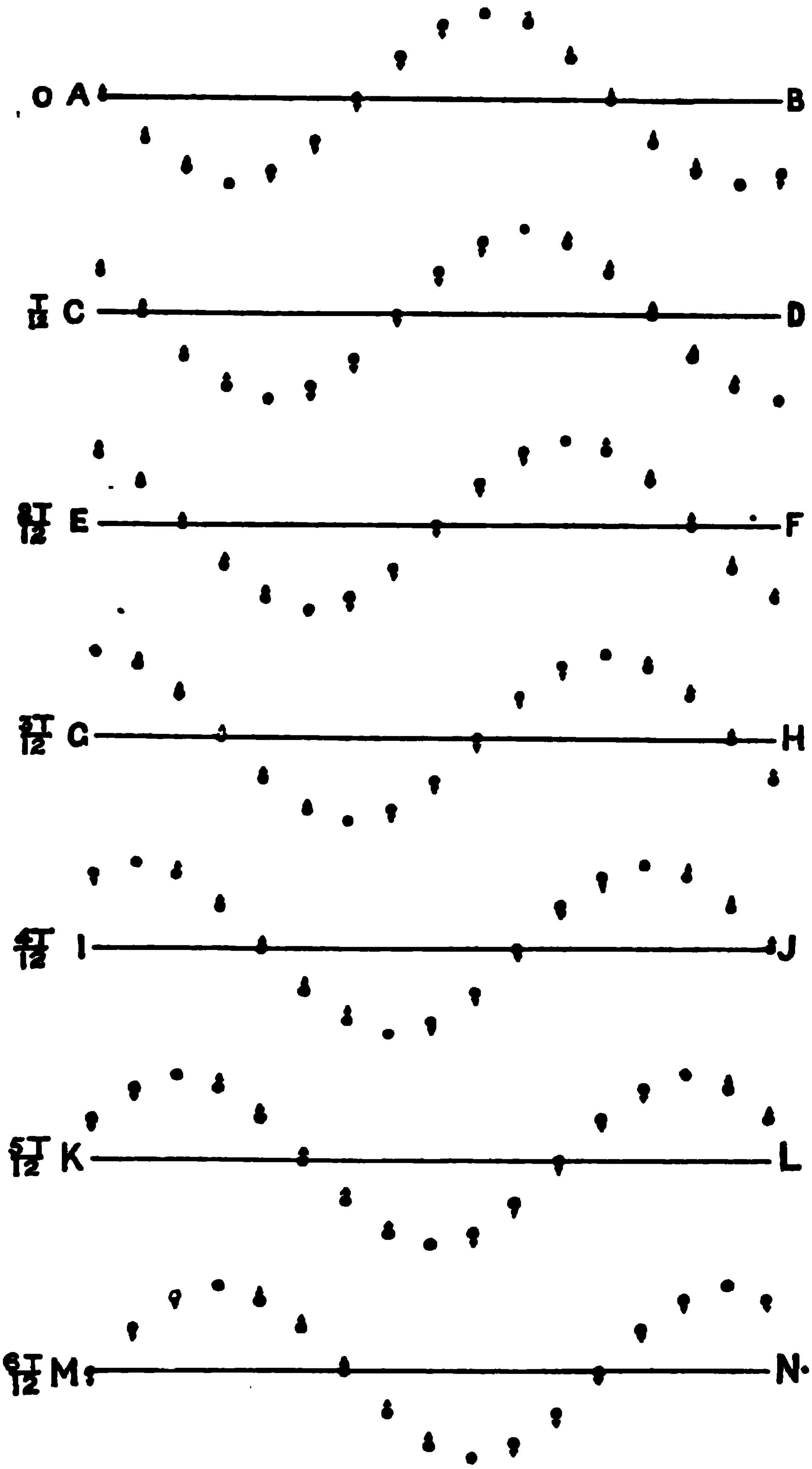


FIG. 215.

amount that particle 1 will be at $2T/12$ from the start, and so on. Hence

the curve drawn through the positions of the particles at any instant will be a harmonic curve (§ 52). Particle 13 will at every moment be in exactly the same state as particle 1, particle 14 as particle 2, and so on; for, as their phases differ by a whole period, they will be equally displaced and *moving in the same direction.*

In Fig. 215 the positions of the particles are shown at successive intervals of $\frac{T}{12}$ up to half a complete vibration from the positions depicted in the first line, the direction of motion at the given instant being indicated by an arrow-head. It will be seen that the curve drawn through the particles can in each case be obtained by displacing the curve for the preceding configuration to the right, and hence, as the motion goes on, the curve connecting the particles appears to move steadily to the right. The distance through which it moves during one complete period of one of the moving particles is equal to the distance between two particles which are moving at every instant in the same direction and are equally displaced on the same side of their mean positions. This distance through which the curve, called in this case a wave, moves during a complete period of one of the moving particles is called the *wave-length* of the motion. The wave-length may also be defined as the distance between one particle and the next one that is displaced from its mean position to the same extent and is moving in the same direction, that is, between two consecutive particles which are in the same phase. Thus in Fig. 214 the wave-length is equal to $\overline{\mathrm{AC}}$ or $\overline{\mathrm{DE}}$.

Although the form of the wave is similar to the harmonic curve, it must be remembered that the harmonic curve represents the successive displacements of a single particle, the abscissæ representing time, while the wave-form curve represents the simultaneous positions of a number of particles, the abscissæ being the distance of the mean positions of the particles measured from some fixed point. However, as all the particles move in exactly the same way, and in one whole wave-length we shall have an example of a particle in every phase of this motion, we may look upon the wave-curve as also showing us what the displacement of each particle will be at different times.

A point on the wave such as F (Fig. 214), at which the particle is at its maximum positive displacement, is called a *crest*, while a point such as D or E, where the displacement has its maximum negative value, is called a *trough*. The positions of the crests and troughs appear to travel towards the right as the motion of the particles continues.

This translatory motion of the wave is not accompanied by the translation of the particles themselves, that is, although each particle moves to and fro along its own little path, yet its mean position during a complete oscillation remains unaltered. We may, therefore, define a wave as a form or disturbance which travels through a medium, and is

due to the parts of the medium performing in succession certain periodic motions about their mean positions.

In the case of wave-motion considered above, the particles all vibrated at right angles to the direction in which the wave moves, and this form of wave-motion is said to be due to *transverse vibrations*. If the motion of each particle takes place in the direction in which the wave moves, then the vibration is said to be *longitudinal*.

At AB (Fig. 216) the undisplaced positions of the particles are shown. If each particle now executes a S.H.M. in the direction AB, the period and amplitude being the same for all, but the phase of each particle being 30° behind that of the preceding particle, then when particle 1 is

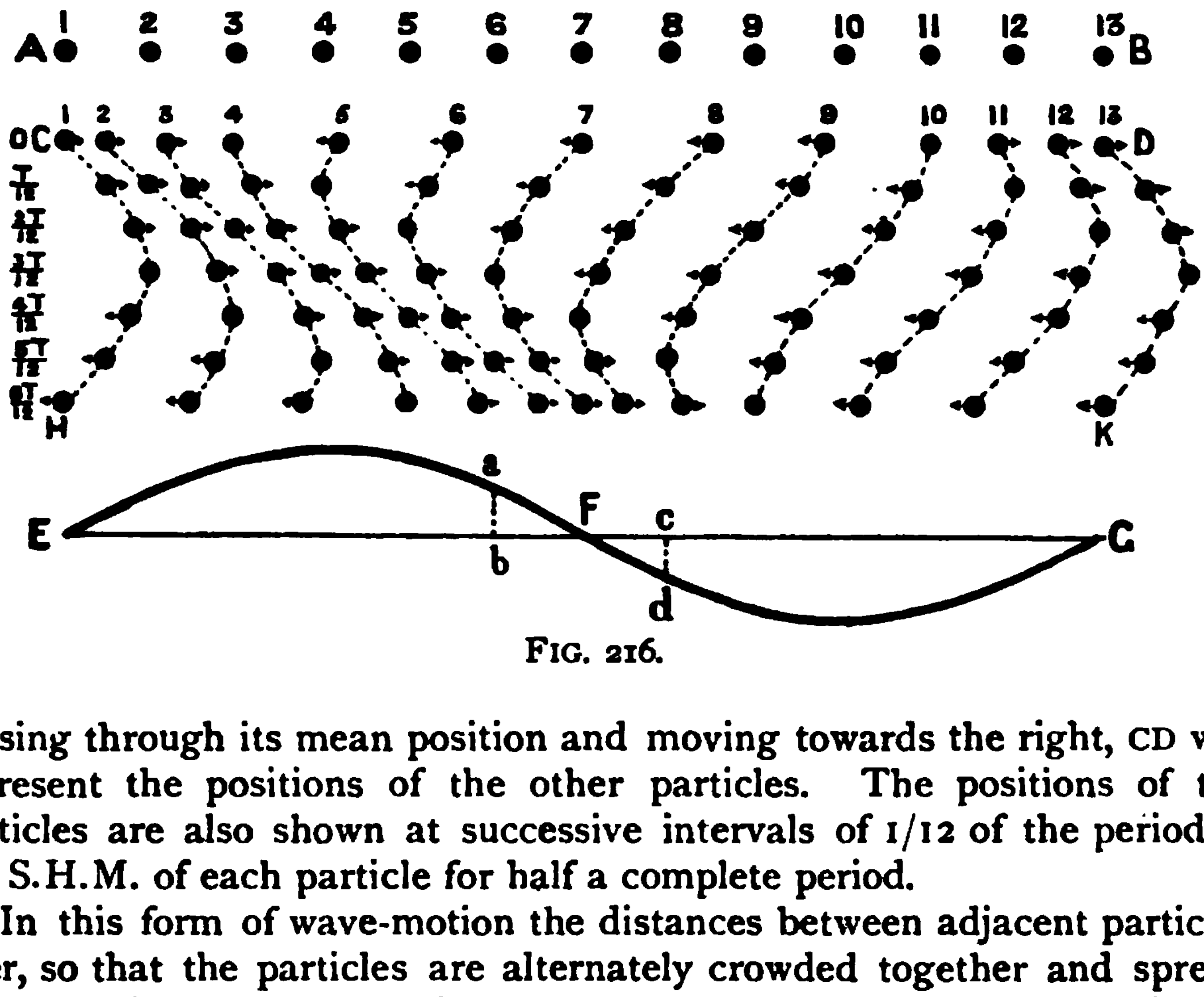


FIG. 216.

passing through its mean position and moving towards the right, CD will represent the positions of the other particles. The positions of the particles are also shown at successive intervals of 1/12 of the period of the S.H.M. of each particle for half a complete period.

In this form of wave-motion the distances between adjacent particles alter, so that the particles are alternately crowded together and spread out. A point where at any instant the crowding together is a maximum is called a condensation, while a point where the distance between adjacent particles is a maximum is called a rarefaction. These play the same parts in longitudinal wave-motion as do the crests and troughs in transverse wave-motion.

The definition of wave-length, given with reference to transverse vibrations, applies also to longitudinal vibrations. The most convenient manner of studying longitudinal vibrations is to employ a curve of which the ordinates indicate the displacements from their mean positions of the different particles at any time. Such a curve is obtained if, at the mean or undisturbed position of each particle, we erect a perpendicular in the positive or negative direction according as the displacement of the

particle is in the positive or negative direction, and having a height equal to the displacement of the particle from its mean position. The curve obtained by joining the extremities of these ordinates is shown at EFG (Fig. 216), the corresponding positions of the disturbed and undisturbed particles being shown at HK and AB. This curve is a harmonic curve, and the points where the curve cuts the axis correspond to the places where the particles are most crowded together, or most spread out. For at F the particle 7 is at its mean position, while the particle 6 is, since the corresponding ordinate *ab* of the curve is positive, displaced to the right by an amount equal to this ordinate, and the particle 8 is to the left of its median position by an amount equal to the ordinate *cd* of the curve; so that the particles are here crowded together. In the same way the particles at E and G are separated to a maximum extent. Hence F corresponds to a condensation, while E and G correspond to rarefactions. The distance between two adjacent rarefactions, such as E and G, or between two condensations, is equal to the wave-length of the wave-motion, while the distance between a rarefaction and the adjacent condensation is equal to half a wave-length.

267. Velocity of Propagation of a Wave—Frequency.—The speed at which the crest or trough in the case of a transverse wave, or the condensation or rarefaction in a longitudinal wave, moves through the medium is called the velocity of propagation of the wave-motion.

While particle 4 in Figs. 214 and 215 is making a complete oscillation, the trough of the wave will travel to the right to particle 16, that is, through a distance equal to the wave-length, λ. In the same way, while particle 1 (Fig. 216) is making a complete oscillation, the condensation will travel from C to D, that is, through a distance equal to the wave-length.

Hence if T is the time each particle takes to complete one oscillation, in this time the wave will move through a distance equal to the wave-length. Thus if v is the velocity of propagation of the wave, we have—

$$v=\lambda/T.$$

Each time that particle 10 (Fig. 214) reaches its maximum positive elongation, a crest will be passing at F, so that the interval between the passage of two successive crests is T. Thus if n is the number of crests which pass F in a second, we have $n=1/T$. The same remark applies to any other particle, whether the motion is transverse or longitudinal, and the quantity n is called the frequency of the waves. Thus

$$v=n\lambda.$$

The velocity with which a group of waves moves into an undisturbed portion of the medium is not necessarily equal to the velocity of the individual waves. Thus in the case of gravitational waves on a liquid, the individual waves travel twice as fast as does the front of the disturbance. Thus if we watch a short train of waves moving into

still water, the waves will appear to move through the group, dying out in front, and fresh waves appearing in the rear of the group. It can be shown that whenever the velocity of the waves varies with the wave-length, the group velocity is different from the wave velocity.

The study of waves being of very great importance in physics—for, as we shall see, sound, light, radiant heat, and many electro-magnetic phenomena are propagated by wave-motions—it will be advisable to spend some time considering this form of motion. It will add to the interest, and also to the clearness, of the study of a wave-motion if we illustrate the various points by reference to some particular form of wave-motion. Now the waves which constitute sound, light, and heat are invisible, and so for the purposes of illustration it will be better to consider the waves which may be produced at the surface of a liquid, for such waves may, with suitable arrangements, be seen by the eye.

268. Waves on the Surface of a Liquid.—In order that a wave may be formed, it is necessary that the successive particles which constitute the medium in which the wave is propagated should each in succession go through a periodic motion. Now when considering the motion of a pendulum (§ 112), we showed that the reason it executes its periodic motion is that, when the bob is displaced, a force acts on the bob tending to bring it back to its position of rest. Hence when dealing with the production of a wave-motion in a medium, we must consider how the force of restitution on the particles of the medium, which is necessary for the production of the periodic motion of these particles, is brought about. Let AB (Fig. 217) represent the plane surface of a liquid at rest. Now suppose by some means we cause the liquid to be heaped up in the form DEF, or scooped out into a hollow FGH, so that the liquid particles are displaced from their positions of rest. Then, owing to the action of gravity, the particles in the portion DEF of the liquid will move back towards the level surface AB, while the particles which have been forced down owing to the production of the hollow FGH will move up. Thus when the particles of a liquid are moved, so that a portion of the surface is displaced either above or below the level of the general surface, owing to gravity a force will act tending to bring the surface back to its undisturbed position. We have therefore the conditions suitable for the production of waves on the surface of the liquid, and the existence of these waves being due to the action of gravity, they are called gravitational waves. The large waves seen on the surface of the sea are well-known examples of gravitational waves. Gravity, as was, however, first pointed out by Lord Kelvin, is not the only cause tending to bring the surface of the liquid

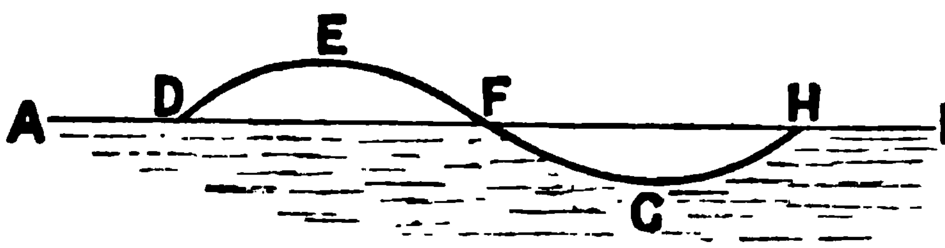


FIG. 217.

back to its undisturbed position. There is a second cause acting, namely, the surface tension (§ 157) of the surface film of the liquid. This surface tension acts as if there were a thin elastic membrane stretched over the surface, and it is evident that the effect of such a stretched membrane will be to tend to flatten down the portion DEF of the disturbed surface of the liquid, and to level up the portion FGH, so that in the surface tension we have also a force of restitution acting on the displaced liquid particles. We have seen in § 158 that the pressure due to the surface tension increases with increase of the curvature of the surface, so that, since the magnitude of the surface tension is small, it is only when we are dealing with waves in which the curvature is very great that we need take account of surface tension.

It can be shown,[1] although to do so would lead us beyond the scope of this book, that if v is the velocity of a wave along the surface of a liquid of which the depth is not less than the wave-length λ, and ρ and T are the density and surface tension of the liquid, then

$$v^2=\frac{g\lambda}{2\pi}+\frac{2\pi T}{\lambda\rho}.$$

From this expression it follows at once that if the wave-length λ is great, the fraction $2\pi T/\lambda\rho$ is small compared to $g\lambda/2\pi$, and hence may be neglected. The fact that λ is great shows that the curvature of the surface must be small, so that this result is what we should expect. On the other hand, if λ is small, then $2\pi T/\lambda\rho$ is great compared to $g\lambda/2\pi$, so that in this case surface tension plays the important part in the propagation of the waves. Such waves, in which the greater part of the force of restitution is due to surface tension, are called capillary waves or ripples.

For waves of wave-length greater than about 4 inches or 10 cm. the term $2\pi T/\lambda\rho$ may be neglected, while for waves of wave-length less than 0.1 inch or 3 mm. the term $g\lambda/2\pi$ may be neglected. For waves having wave-lengths between these two limits, we have to take into account both the effect of gravity and of surface tension.

Since the velocity due to gravity alone increases as λ *increases*, and that due to surface tension alone increases as λ *decreases*, it follows that there must be a certain wave-length for which the velocity is a minimum. For wave-lengths less than this critical value the surface tension has the predominating influence, and therefore the velocity increases as λ decreases, as shown by the left-hand branch of the curve (Fig. 218),[2] which

[1] The case where gravity *alone* is supposed to act is considered in § 277.

[2] In order to show a considerable range, instead of taking equal length along the axes to correspond with equal increments in the wave-length and the velocity respectively, in the figure equal lengths along the axes correspond to equal increments in the logarithms of these quantities. The scales have, however, been numbered so that we read off the wave-length and velocity direct. For an account of this method of plotting curves, a paper by Prof. C. V. Boys, in *Nature* for July 18, 1895, may be consulted.

gives the velocity of waves of different wave-lengths in water. For wave-lengths greater than the critical value, gravity plays the important part, and the velocity increases as λ increases, as shown by the right-

Fig. 218.

hand branch of the curve. For water the minimum velocity is 23 cm. per second, or 9 inches per second.

269. Gravitational Waves.—In the case of waves for which the wave-length λ is so great that we may neglect the effect of surface tension, we have

$$v^2 = \frac{g\lambda}{2\pi}.$$

It will be observed that the density of the liquid is not involved in the expression for the velocity. The reason for this is the same as that which explains why it is that the period of a pendulum is independent of the mass of the bob, namely, that although the mass of the liquid to be moved is proportional to the density, yet, since the force of restitution is also proportional to the density, for it is the weight of the raised portion of the liquid, the ratio of the force of restitution to the mass to be moved is the same for all liquids, and therefore the velocity of the waves is the same.

If the depth of the liquid is considerably less than the wave-length, the velocity is less than that given above, and is given by

$$v^2 = gd,$$

where d is the depth of the liquid. One effect of this decreased velocity in shallow water is to make the waves in the neighbourhood of a shelving beach always move in a direction perpendicular to the shore, although at some distance out to sea they may be moving in quite a different direction. The reason is that when a wave which is moving in a direction inclined to the shore-line reaches shallow water, the end of the wave which first reaches the shallow moves more slowly than the parts which are still moving in deep water. Thus the wave gradually wheels round till it becomes nearly parallel to the shore.

In the case of a wave in deep water, the individual particles of the water describe circles in vertical planes as illustrated in Fig. 219, where the form of the wave is shown at two instants corresponding to an interval of one-twelfth of the periodic time. Thus when a particle is on the crest, A, of the wave, it is moving in the direction in which the wave is moving, while when it is in the trough, B, it is moving in the opposite direction to that in which the wave is moving. As we go down from the surface the particles still move in circles, but the radii grow

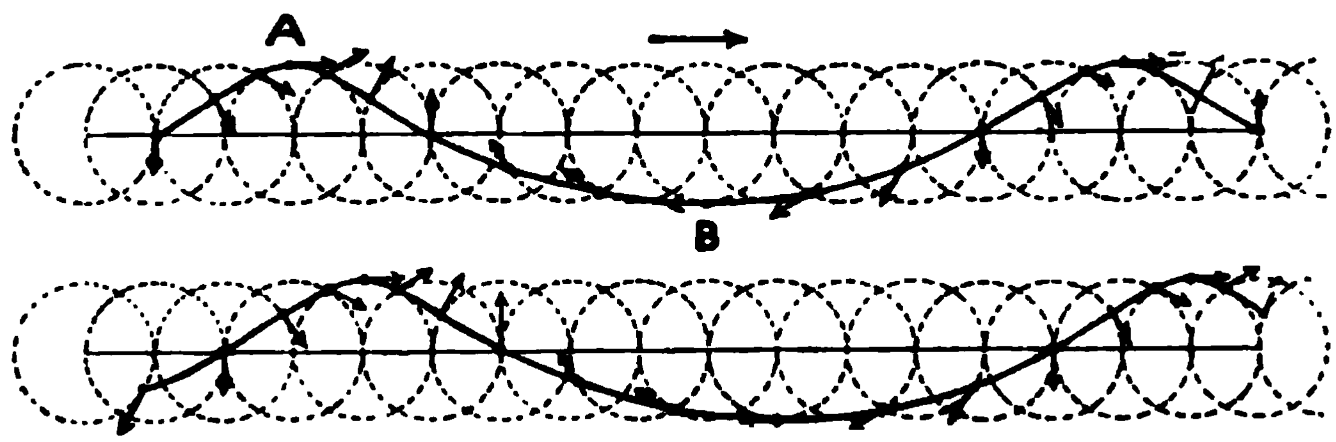


FIG. 219.

smaller and smaller, till at a depth equal to the wave-length the radius of the circles is only about 1/500th of what it is at the surface. In shallow water the paths of the individual particles are ellipses with their major axes horizontal. In this case the horizontal axes of the ellipses are approximately the same for particles at all depths. The vertical axes, however, decrease with the depth, till at the bottom they vanish, and the particles move backwards and forwards along straight lines.

When the height of the crest of a wave above the undisturbed level of the water is equal to the depth of the undisturbed water at the point, the particles at the crest will be moving forward with the same velocity as the wave, and the wave will be unstable and "break." As a wave comes into shallow water the wave-length decreases, for the velocity decreases as the depth of water decreases, and the frequency (n) must remain the same, that is, the number of waves which pass a given point in one second, and $v=n\lambda$. The effect of this shortening of the wave-length is to make the amplitude of the waves greater. This goes on, till finally the unstable condition is reached, and the wave breaks.

270. Capillary Waves.—In the case of waves of which the wave-length is less than 4 mm., we have

$$v^2=\frac{2\pi T}{\lambda\rho}.$$

Here both T and ρ depend on the nature of the liquid, so that the

velocity of capillary waves is different in different liquids. If n is the frequency of the waves,

$$v = n\lambda,$$

and
$$n^2\lambda^2 = 2\pi T/\lambda\rho,$$
or
$$T = n^2\lambda^3\rho/2\pi.$$

Thus if the frequency n is known, and we measure the wave-length, we can calculate the surface tension T. Lord Rayleigh has used this method for measuring the surface tension. The waves were produced by a fine style attached to the prong of a tuning-fork which dipped into the liquid. Thus the frequency of the waves was equal to the frequency of the fork, and was therefore known.

271. Interference of Waves.—If we have two systems of waves passing over the same surface of water, each will produce the same effect as if it were alone present, so that the actual displacement of any surface particle of the water at a given instant is the algebraical sum of the displacement it would have, at that instant, due to each set of waves separately. The resultant motion is thus obtained by compounding the two separate wave-motions, just as in § 54 we obtained the resultant motion of a point when moving with two simple harmonic motions.

This separate existence of two sets of waves is one of everyday observation, when two stones are thrown into still water. Each stone will produce a set of waves which travel out in ever-widening circles, and the circular waves due to one stone will pass unchanged through the waves due to the other.

Suppose we have a style attached to one of the prongs of a tuning-fork dipping into a vessel containing liquid, say mercury. When the fork is in motion the style will produce a system of waves which will move out in circles from the point where the style enters the mercury. The radius of each of the circular waves will increase at the rate given by

$$v = n\lambda,$$

where n is the frequency of the fork, and λ the wave-length as given by the equation

$$\lambda^3 = 2\pi T/n^2\rho.$$

Let the position of the waves at a given instant be as represented by the circles, with A as centre, in Fig. 220, where the heavy full lines represent the crests, and the heavy dotted lines the hollows. The waves in only half the circumference are drawn, in order to save space. Now, suppose there is a second style attached to the same prong touching the mercury surface at B, so that whenever a crest starts from A an equal crest will start from B, and so on. The position of the waves due to B alone are shown in the figure by the light lines. The wave-length and velocity of the waves starting from B will be the same as those starting from A. In order to obtain the actual condition of the surface due to the combined action of the two sets of waves, we have at every point to add

together the displacements due to each set. Thus at the point C we have a crest of each set of waves, so that the upward displacement here is twice the maximum displacement due to either set separately. At D, in

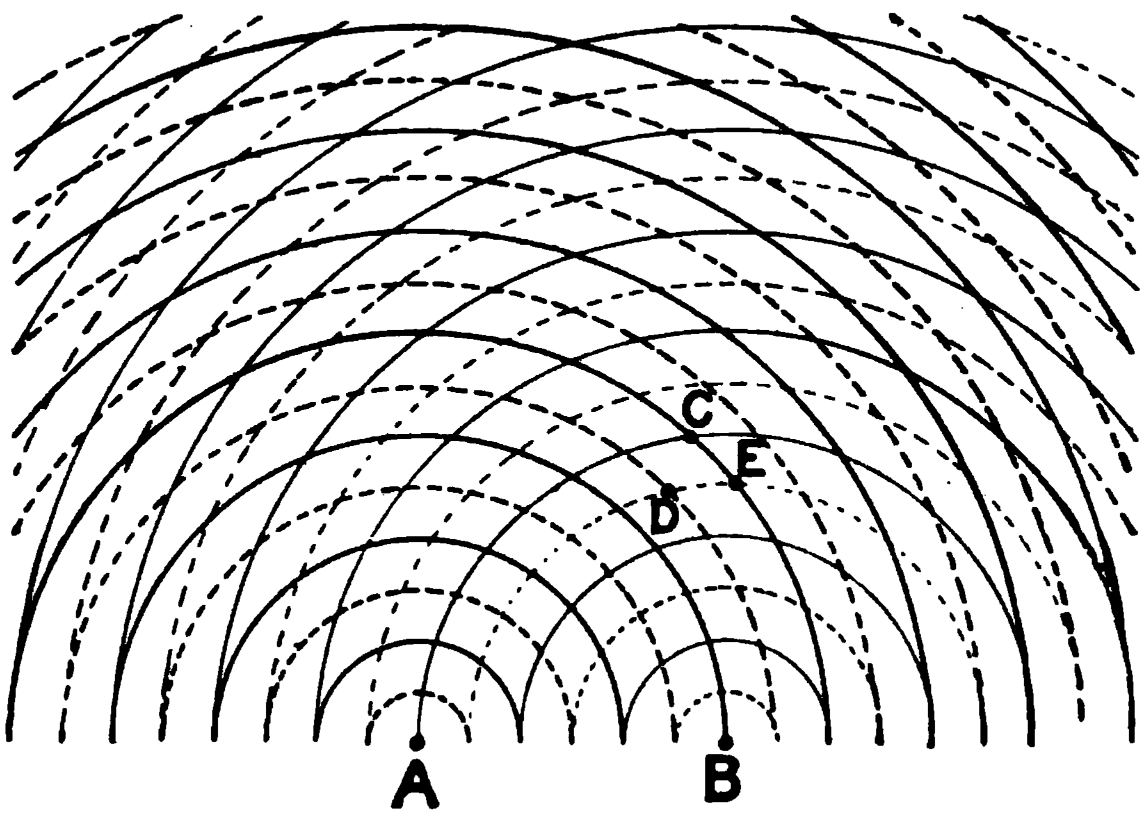


FIG. 220.

the same way, we have a trough due to each set, and the downward displacement is double that due to either set of waves alone. At E, however, we have a crest due to the waves starting from A, and a trough due to the waves starting from B. The result is that the particle at E is

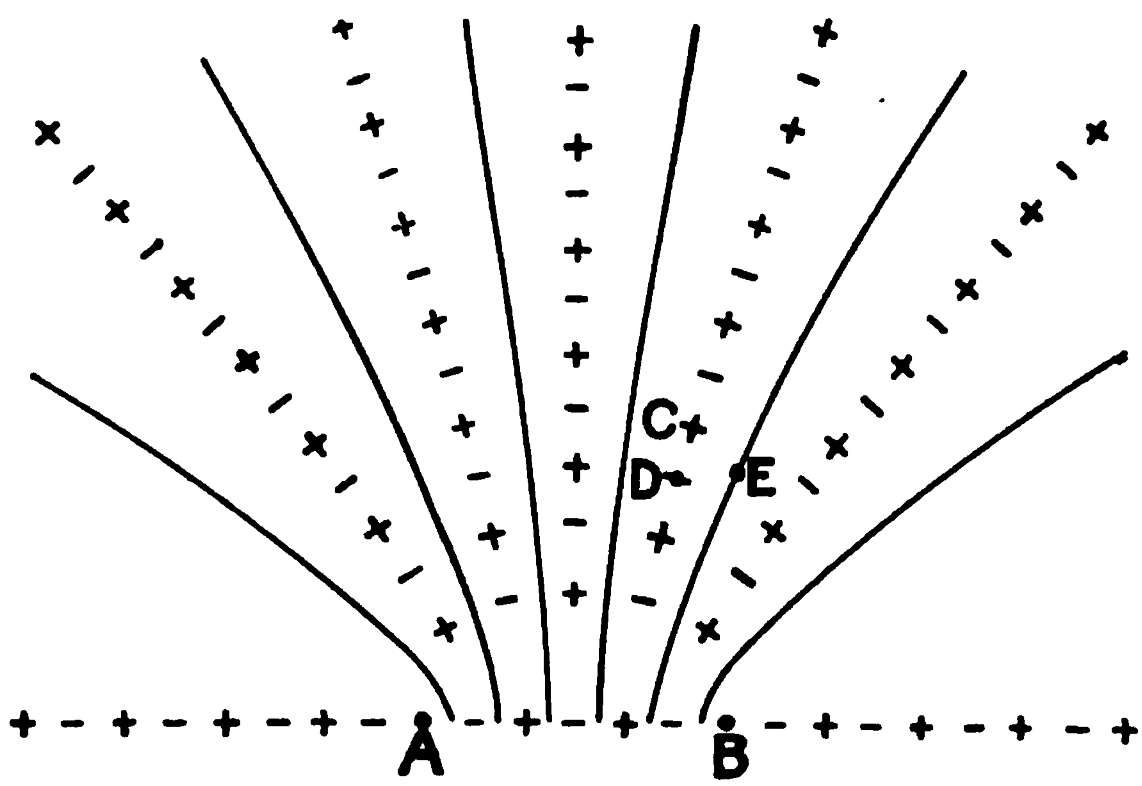


FIG. 221.

undisplaced, for the upward displacement due to one set of waves is just neutralised by the downward displacement due to the other. In Fig. 221 lines are drawn through the points which are undisturbed, while a + sign

marks the points where there is maximum upward displacement, and a − sign the points where there is a maximum downward displacement.

Next, suppose we again draw the figure for a time equal to half a period later. Each of the waves will have travelled out through a distance equal to half a wave-length, and so a crest will now occupy the position previously occupied by a trough, and *vice versa*. Thus the conditions are still represented by Fig. 220, where, however, the dotted lines now represent the crests and the full lines the troughs. The point C is now at a trough, and the point D at the crest of the disturbance due to the two sets of waves. The point E is, however, still at rest, for now it is at the trough of a wave from A, and at a crest of a wave from B. It will also be found that all parts of the liquid surface along the lines drawn in Fig. 221 are still at rest. In this way it can be shown that, owing to the joint action of the two sets of waves, we have certain portions of the mercury surface which are permanently at rest, although, if either set of waves acted alone, these parts would be disturbed by the passage of waves. This phenomenon, of a state of rest being produced by the combined action of two sets of waves, is called *interference*, and we shall find that it plays a very prominent part in many natural phenomena.

FIG. 222.

That we actually do get these lines of no disturbance in the case of capillary waves can be seen by eye, for, although the individual waves cannot be distinguished on account of their rapid motion, yet the undisturbed portions of the surface appear brighter than the rest, and we see a pattern similar to Fig. 222. If, instead of looking at the surface, we take an instan-

FIG. 223.

taneous photograph, then, as shown in Fig. 223, we see not only the lines of no disturbance but also the waves between these lines, which are due to the combination of the two sets of waves.

272. Wave-Front—Ray.—Suppose a medium is disturbed by the passage of a system of waves and we draw a curve, if the waves extend only in two dimensions as is the case with water waves, or a surface, if the waves extend in three dimensions as is the case with sound waves in free air, such that it everywhere passes through portions of the medium which are in the same phase (§ 50) of their vibratory motion, then such a line or surface is called a *wave-front.* If in the case of water of uniform depth a disturbance is produced at a point *a*, then the waves will travel out in circles from *a* as centre, and all the water particles on the circumference of a circle described with *a* as centre will be in the same phase, that is, they will be at equal distances from, and on the same side of, their undisturbed positions. Thus any such circle will be a wave-front.

A wave in which the wave-front is either a straight line or a plane is called a plane wave, while one in which the wave-front is a sphere is called a spherical wave.

In the case of a disturbance produced at a point, it is evident that at any point of one of the circular wave-fronts, the wave is moving along a radius of the circle, that is, at right angles to the tangent to the wave-front at the point considered. This result is quite general, so long as the medium in which the wave is propagated is isotropic, the direction of motion of the wave being always at right angles to the wave-front.

In many problems we have only to do with the direction of motion of the waves which are being considered, and a line drawn, so as everywhere to indicate the direction of motion of the waves, is called a *ray.* Thus the rays are everywhere at right angles to the wave-fronts. Since the waves must start from the centre of disturbance—and if this is a point, the wave-fronts are circles with this point as centre—the rays must all pass through the centre of disturbance. If the medium is isotropic, the rays will be straight lines; if, however, the medium is not isotropic, then the rays may be curved.

If from every point of a small portion of the wave-front we draw a ray, the resultant system of lines is called a *pencil of rays.*

273. Huyghens's Construction.—In the place of two centres of disturbance, such as those considered in § 272, suppose we have a number placed in a line AB (Fig. 224), where, for simplicity, the position of the crests only are shown for three waves due to each centre. It will be observed that along the lines which touch the circular waves due to all the centres of disturbance, the waves all work together to form a crest, while at other places we have a crest due to one set coinciding with a trough due to others. The result is that, except along the tangents, there will be very little disturbance. Of course half-way between the

crests we should, if we drew the circles to represent the positions of the troughs, have a resultant trough which would also be a straight line parallel to AB. It will be understood how, if the number of centres of disturbance between A and B was made very great, the resultant effect would be a series of waves which would form straight lines parallel to AB, and would move out over the surface of the liquid in a direction at right angles to AB. We have thus arrived at an explanation of the formation of a plane wave, that is, one in which the crests and troughs are straight lines, and may look upon it as produced by the action of an

A B

FIG. 224.

infinite number of disturbing centres placed along the straight line AB. Such a plane wave is obtained in practice if, in place of a style, we attach a flat glass plate to the prong of a tuning-fork, so that the edge dips into the mercury. The construction for finding the position of one of the crests, that is, of the wave-front, at a time t by means of the tangent to a series of circles, the radius of each of which is equal to the space passed over by the wave in a time t, is due to Huyghens, and is known as Huyghens's construction for the wave-front. The centres of disturbance need not lie on a straight line. Thus suppose they lie on a circle ABC (Fig. 225), of which the centre is D and radius R. If now from all points along the circumference of this circle we describe circles of radius r, where r is the distance the wave will travel through in a time t, we shall get the wave-front at a time t by drawing a line touching all these circles. This tangent is evidently a circle of radius $R+r$, having its centre at D. Now if a disturbance were produced at D, we know that in a time R/v, where

F

FIG. 225.

v is the wave velocity, the wave-front would be the circle ABC, while at a time t later it would be the circle EFG, so that if we like we may look on each point on the wave-front ABC as being a centre of disturbance, and consider that it is due to the combined action of these secondary centres of disturbance that the wave-front EFG is produced, and not directly by the disturbance at D. We shall find Huyghens's construction of considerable use in future for finding the position of the wave-front at any time subsequent to that at which the wave-front has some given position.

274. Reflection of Waves.—When a wave meets an obstacle it is not in general annihilated, but its direction of propagation, and therefore also its wave-front, becomes altered. This phenomenon is called reflection. Thus suppose AB (Fig. 226) represents a wall limiting a stretch of water, and CED represents a wave moving in the direction of the arrow. As this wave moves forward, first the end C and then each part in succession will meet the wall and will be reflected, so that, in addition to the portion E′D′ of the wave moving in the original direction, there will be a reflected portion C′E′ moving in the direction shown by the arrow; while, finally, when the whole of the wave has been reflected we shall have a wave C″E″D″, which will move off in the direction indicated by the arrow. In order to find the connection between the inclination of the incident wave-front to the reflecting surface and that of the reflected wave-front, consider the incident wave CL (Fig. 227), of which the end C has just reached the reflecting surface AB. If there had been no reflecting surface, then, when the other end of the wave L reached G, the wave would have occupied the position HG. Let us start reckoning time from the instant when the wave reaches C, that is, from the instant when the wave is in the position CL, and suppose that the time taken by the point L on the wave to reach G is t, so that if v is the velocity with which the wave moves,

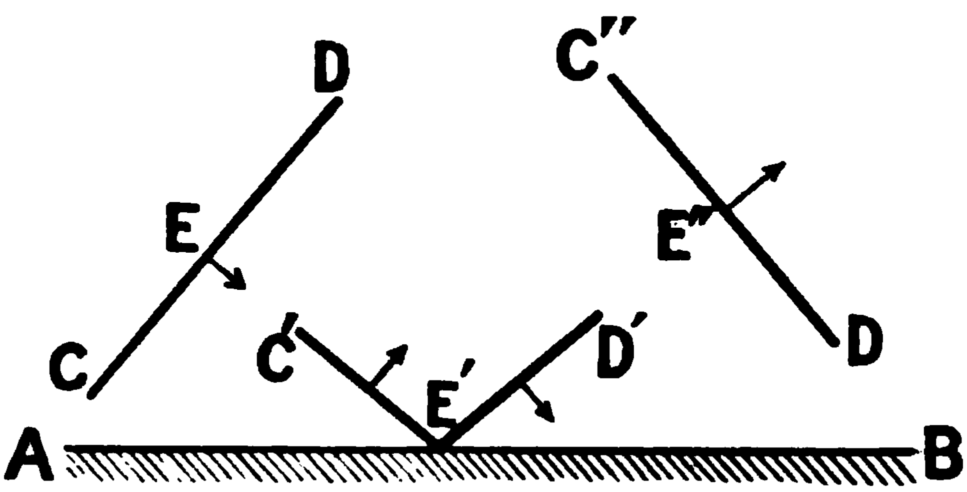


FIG. 226.

$$t = \frac{\overline{LG}}{v}.$$

Now each of the lines CH, DD‴, EE‴, FF‴, are equal to LG, and represent the distance moved through by the wave in the time t. When the end C of the wave reaches the reflecting wall, we may consider that C becomes a centre of disturbance, so that in a time t the wave produced by this centre will form a circle of which the radius r is given by $r = vt$, that is,

$r = \overline{LG} = \overline{CH}$. Hence, if a circle is described with C as centre and radius $\overline{CH}$, it will represent the position of the reflected wave from the point C at the time t. Next, consider some other point D on the wave. It will have to move over the distance DD′ before it meets the reflector, and when it does so, we may regard it as becoming a centre of disturbance. Since, if it were not for the reflection, in the time t, the point D on the

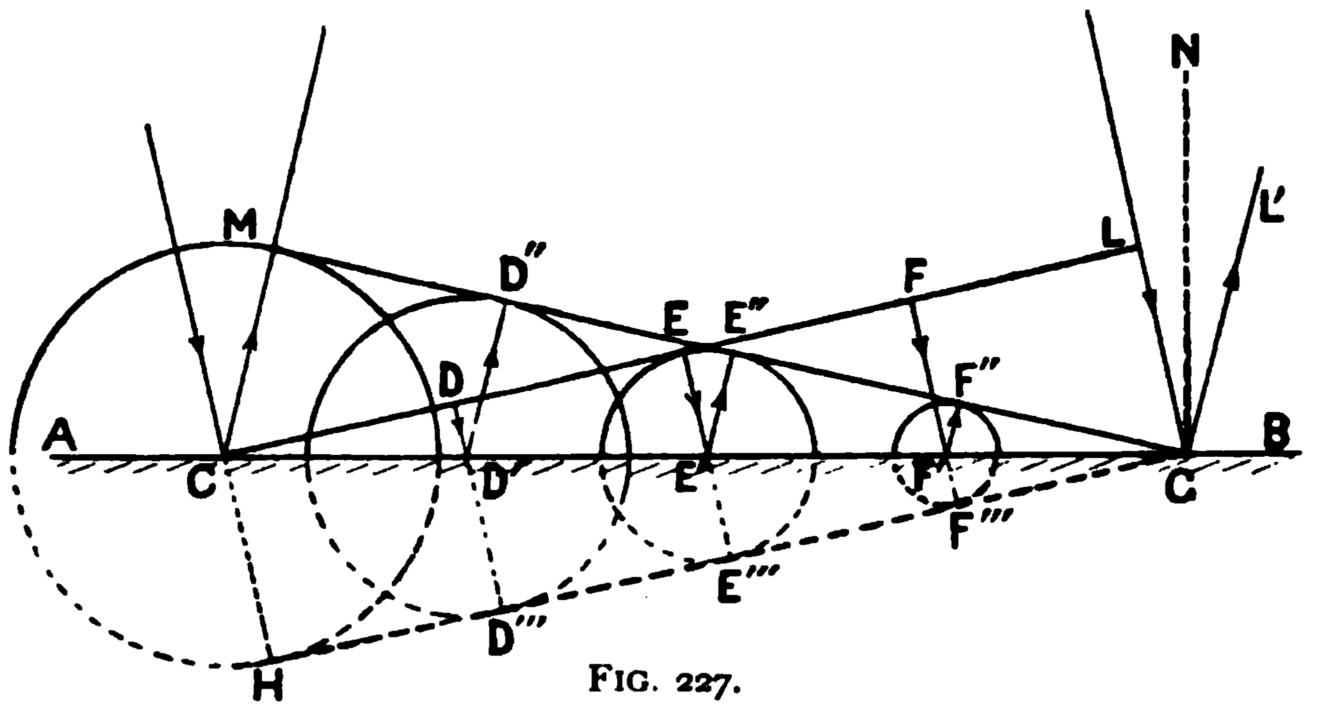


FIG. 227.

wave would travel to D‴, it is clear that in the time t the wave will travel up to D′, *i.e.* through the distance $\overline{DD'}$, and then have time to travel through a further distance equal to $\overline{D'D'''}$. Therefore, if a circle is described round D′ as centre, with $\overline{D'D'''}$ as radius, this circle will represent the position, at the time t, of the reflected disturbance produced by the point D on the wave.

Proceeding in this way, we can show that the reflected waves formed by each point of the incident wave meeting the reflecting surface touch a straight line MG, for all the centres of the circular waves lie on the straight line AB, and the circles which show the position of the reflected waves at the time t all touch the straight line HG. Hence by Huyghens's principle the line MG represents the position of the reflected wave-front at the time t after the instant when it occupied the position CL.

The two triangles CLG, GCM are equal, for $\overline{LG} = \overline{CM}$, and $\overline{CG}$ is common, and the angles CLG, GMC are each right angles. Hence the angle LCG, made by the incident wave-front with the reflecting surface AB, is equal to the angle MGC, made by the reflected wave-front with the reflecting surface.

The line LG shows the direction in which the incident wave moves, and is the incident ray, while the line GL′ shows the direction in which the reflected wave moves, and is the reflected ray, for these lines are perpendicular to their respective wave-fronts. Since the wave-fronts make equal angles with AB, these perpendiculars must also make equal angles with AB, so that the angle LGC is equal to the angle L′GB. Hence,

if GN is the normal to AB at G, that is, a line drawn through G perpendicular to the reflecting surface, the angle LGN must be equal to the angle L′GN. Now the angle LGN, between the direction of motion of the incident wave and the normal, is called the angle of incidence, while the angle L′GN, between the direction of motion of the reflected wave and the normal, is called the angle of reflection. Hence we see that when a wave is reflected, the angle of incidence is equal to the angle of reflection.

275. Stationary Waves.—In the last section we considered the reflection of a single wave, so that at no time did the incident and reflected waves affect the same particles simultaneously. If, however, we are dealing not with a single wave but with a train of waves, then a given point in the liquid may be affected by one of the incident waves, and at the same time by one of the preceding waves which is returning after reflection. The result of this simultaneous action of two sets of waves, the incident and the reflected, is that interference may take place. We shall only consider the simplest case, namely, that when the direction of the incident waves, and therefore also that of the reflected waves, is at right angles to the reflecting surface.

Let AB (Fig. 228) represent a vertical section of the reflecting surface, and let the waved line CD represent a section through the incident waves. Each of these waves, as it reaches the obstacle, will be reflected, so that we shall have a train of reflected waves travelling away from AB, which for clearness are shown separate at EF.

Since it is evident that when a crest of the incident waves reaches the obstacle a crest will be produced on the reflected wave, and as at the instant for which the figure is drawn a crest on the incident wave is at C, we must have a crest on the reflected wave at E. Owing to the combined action of the incident and reflected waves, the form taken by the water surface is shown at GH, in which the displacement at any point is the sum of the displacements due to the two systems of waves separately.

If τ is the period of the waves, then in a time $\tau/4$ the incident waves will have moved through a distance equal to a quarter of the wave-length, λ, to the left, for the wave travels over a space equal to the wave-length during the period; also the reflected waves will have travelled through a distance $\lambda/4$ to the right, as shown at C′D′ and E′F′. Under the combined action of the two sets of waves the whole surface will be momentarily in its position of rest, as shown at G′H′, for it will be noticed that the effect of the reflected waves is just to neutralise the displacement due to the incident waves. Similarly the actual state of the water-surface, at times $\tau/2$ and $3\tau/4$, is shown at G″H″ and G‴H‴. If these curves are examined, it will be seen that there are certain points, N_1, N_2, N_3, &c., on the surface of the liquid which remain permanently at rest, owing to the interference between the incident and reflected waves. These points are called *nodes*. Half-way between each node the water surface swings up and down to a maximum extent, and these points are called *loops*

or *antinodes.* The portions of the surface between the nodes and loops move up and down, the amplitude of the movement gradually decreasing from the loop to the node. Thus at one instant we have a series of alternate crests and troughs at the loops, then the surface flattens out, and immediately after a series of troughs and crests appear at the

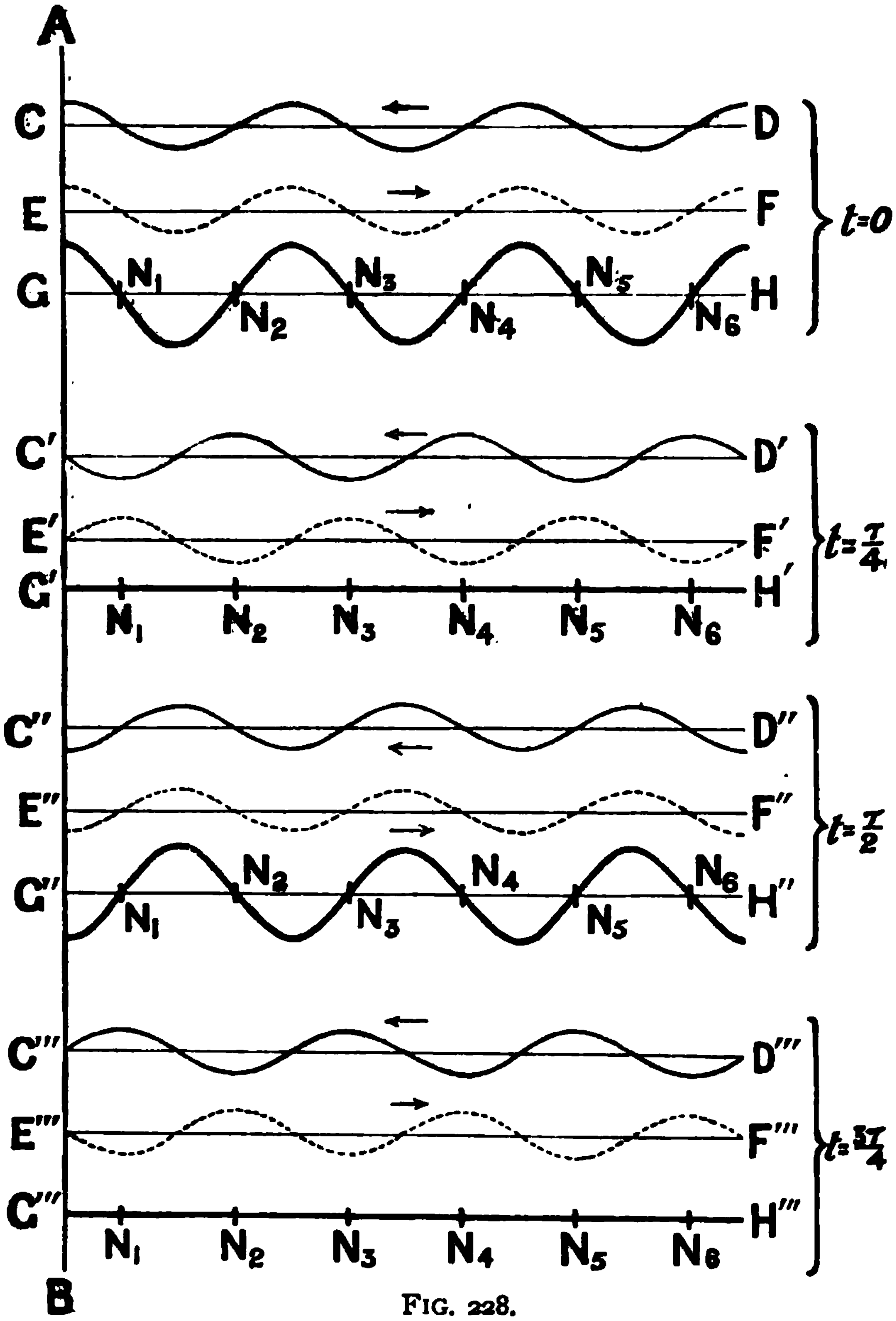


FIG. 228.

loops, and so on; and the character of the disturbance is quite different from ordinary waves, for there is no progressive movement of the crests and troughs. These waves which retain their position unaltered are called stationary waves, and, as we shall find later, they play an important part in many phenomena which involve wave-motion.

It is immediately evident, from Fig. 228, that the distance between consecutive nodes is equal to half the wave-length. Hence if we measure the distance between the nodes, and know the frequency of the waves, we can calculate the velocity with which they travel.

The nodes are points at which the disturbance due to the reflected waves is *always* equal and *opposite* to that due to the incident waves. The loops, on the other hand, are points where the disturbance due to the reflected is equal to, and in the *same* direction as, that due to the direct waves. A consideration of Fig. 228 will show that the portions of the medium on opposite sides of a node are always moving in opposite directions, or are displaced on opposite sides of their positions of rest.

276*. Velocity of Propagation of a Transverse Wave along a Stretched String.—If a string is stretched and then a hump is produced, say by striking the string a sharp blow in a direction at right angles to its length, this hump will travel along the string to the far end, where it will be reflected and will then travel back. This hump is a transverse wave, for each of the particles composing the string moves forwards and backwards along a line at right angles to the string, and the velocity with which it travels depends on the tension T with which the string is stretched, and on the mass m of unit length of the string.

Suppose ABCDE (Fig. 229) to represent the string along which the transverse wave BCD is travelling from left to right with a velocity v. If now the string were made to move from right to left, that is, in the

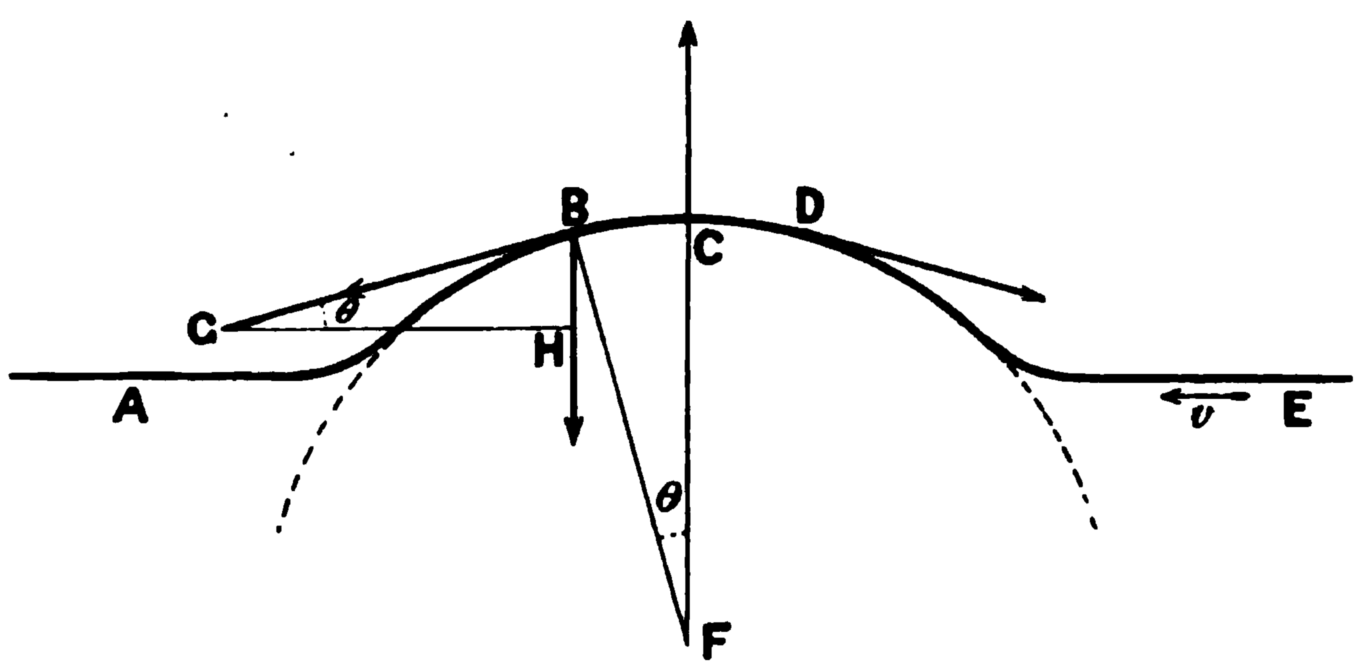


FIG. 229.

opposite direction to that in which the wave is moving, with a velocity v, the wave would remain in one position,[1] for it is moving along the cord with a velocity v, and the cord is itself moving with a velocity $-v$.

[1] We are here supposing that the form of the wave remains unaltered, and that it moves with a constant velocity. A wave which fulfils these conditions is said to be a wave of permanent type. The investigations in the following three sections also only apply to waves of permanent type.

It is an experimental result that the speed with which the wave moves along the cord does not depend on the shape of the hump. Let us therefore assume that, at any rate, the portion near the crest forms an arc of a circle, of which F is the centre. Consider a small portion BCD of the cord at the crest of length δs. The mass of this portion of the cord is $m\delta s$, and since it is moving in a circle with a speed v, it will exert a centrifugal force equal to

$$\frac{m \, . \, \delta s \, . \, v^2}{R}$$

(§ 42), where R is the radius of the circle BCD. This force will act vertically upwards along the radius FC, and therefore, since this element of the cord remains in equilibrium, that is, does not move away upwards under this force, it must be acted upon by an equal and opposite force. Now, the only force which acts upon the element of the cord is the tension of the cord acting at its two ends B and D, which tension acts in the direction of the cord at B and D, that is, along the tangent to the circle at these points. Consider the tension acting at B; it consists of a force T acting in the direction of the tangent BG. We may resolve this force into a vertical component, $\overrightarrow{\mathrm{BH}}$, and a horizontal component, $\overrightarrow{\mathrm{HG}}$ (§ 67). If θ is the angle between the radius FB and the radius FC, since BH is parallel to FC and the tangent BG is at right angles to BF, the angle BGH is equal to θ. Hence the vertical component of the tension $\overrightarrow{\mathrm{BH}}$ is $T \sin \theta$. In the same way the vertical component of the tension at D is also $T \sin \theta$, so that the total vertical component of the tensions acting on the small element BCD of the cord is

$$2T \sin \theta.$$

Now, θ being the circular measure of the angle BFC, the length of the arc BC is $R\theta$ (§ 14). Also, as by supposition the length of the arc BC is very small, so that the angle θ is also very small, the sine of the angle θ is equal to θ (§ 14). Hence the vertical component of the tension is

$$2T \sin \frac{\delta s}{2R} = \frac{T \, . \, \delta s}{R}.$$

Equating this vertical component to the centrifugal force acting on the element of the string, we get

$$\frac{m \, . \, \delta s \, . \, v^2}{R} = \frac{T \, . \, \delta s}{R},$$

or

$$v^2 = \frac{T}{m},$$

or

$$v = \sqrt{\frac{T}{m}},$$

that is, the velocity of the wave is equal to the square root of the tension

of the string, expressed, of course, in absolute force units (§ 61), divided by the mass of unit length.

277*. The Velocity of Gravitational Waves in Deep Water.— In deep water the paths of the individual particles of the water during the passage of a wave are circles, as shown in Fig. 219, so that at the crest B of a wave ABCDE (Fig. 230) which is moving from left to right, the water particles are also moving to the right, while at the trough D they are moving to the left with the same velocity.

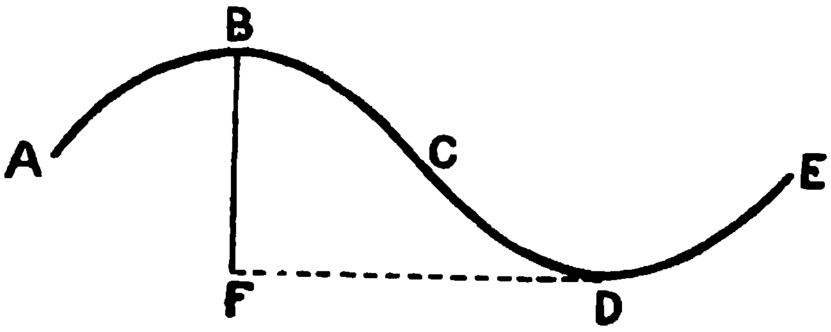


FIG. 230.

Let the velocity of the wave be v, and suppose, as in the last section, that we bring the wave to rest by imparting a velocity $-v$ to the water, that is, in the direction from right to left. Let a be the radius of the circle in which a surface particle moves, then the speed of the particle due to its motion in this circle is $2\pi a/\tau$, where τ is the period of the wave. Thus when the wave is brought to rest by imparting a velocity $-v$ to the water, we have, if we call the velocity from left to right positive, that the velocity of a particle at B is $2\pi a/\tau - v$, and that of a particle at D is $-2\pi a/\tau - v$. Suppose we consider a small volume element of the fluid which, for simplicity, we may take as of unit mass. The kinetic energy of this element when it is at D is

$$\tfrac{1}{2}(2\pi a/\tau + v)^2,$$

while when it reaches B its kinetic energy is

$$\tfrac{1}{2}(2\pi a/\tau - v)^2.$$

Thus the kinetic energy lost between D and B is

$$\frac{4\pi av}{\tau}.$$

Now between the points D and B the element has been raised through a height FB or $2a$, and has therefore gained potential energy of amount $2ga$.[1] Since the gain of potential energy must be equal to the loss of kinetic energy, we have

$$\frac{4\pi av}{\tau} = 2ga,$$

or

$$v = \frac{g\tau}{2\pi}.$$

[1] The surface of the liquid must all be at the same pressure, so that no work is done against hydrostatic pressure while the volume element passes from D to B.

Now if λ is the wave-length, this is the distance passed through by the wave in the time τ, so that

$$v=\lambda/\tau \text{ or } \tau=\lambda v.$$

Hence, substituting this value for τ,

$$v=\frac{g\lambda}{2\pi v},$$

or

$$v^2=\frac{g\lambda}{2\pi}.$$

$$\therefore \quad v=\sqrt{\frac{g\lambda}{2\pi}}.$$

278*. The Velocity of Gravitational Waves in Shallow Water. —Let ABDE (Fig. 231) represent a section of a wave in shallow water, GH being the bottom of the water, and AE the undisturbed surface. Let the velocity of the wave be v, and, as before, suppose the wave brought to rest by giving the water a velocity $-v$. Let d be the depth ML of the water, and a be the height of the crest above the general level or the depth of the trough below this level. Then DK $=d-a$ and BL $=d+a$.

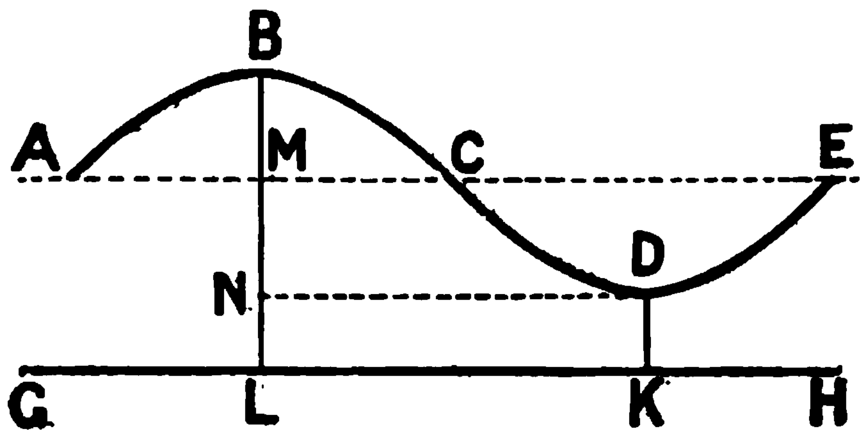


FIG. 231.

Now, in the case of waves in shallow water, the horizontal component of the motion of all the water particles in the same vertical line is the same. Let the velocity with which the particles along the line BL would be moving forward, if the water were at rest, be u, and that with which the particles along the line DK would be moving back, be $-u$.

When the water is moving as a whole with a velocity $-v$, the actual velocity of the water particles at BL is $u-v$, while that of the particles at DK is $-u-v$. Hence, if we consider a slice of the wave of unit breadth perpendicular to the paper, the volume of water which passes through BL in a second is

$$(u-v)(d+a).$$

Also the quantity of water which passes through DK in a second is

$$-(u+v)(d-a).$$

Now as the wave remains in a fixed position, so that B is always a crest and D a trough, the volume of water between BL and DK must remain

constant. Hence just as much water must enter through DK in each second as leaves through BL. Thus—

$$(u-v)(d+a)=-(u+v)(d-a),$$

or

$$a=ud \quad . \quad . \quad . \quad (1).$$

Next, if we consider the motion of a small element of volume which moves from D to B, we shall have, as in the last section, by equating the decrease in kinetic energy to the increase in potential energy,

$$2uv=2ga \quad . \quad . \quad . \quad (2).$$

From (1) and (2)—

$$v=\frac{d}{a}.\frac{ga}{v},$$

or

$$v=\sqrt{dg}.$$

Since the pressure at the surface of the liquid is everywhere the same, the pressure at D and B is the same, and so no work is done against an opposing hydrostatic pressure as the element of volume of the fluid moves along the surface from D to B. If, however, we considered an element of volume which moves from K to L along the bottom, no work will be done against gravity, for the element will be at the same height at L and K. The hydrostatic pressure at L, due to the head $d+a$, is greater than that at K, due to the head $d-a$, and so work will have been done in moving the element against this opposing hydrostatic pressure, and this work will be the equivalent of the decrease in kinetic energy. In the case of a particle between the surface and the bottom, work will be done against both gravity and hydrostatic pressure.

The velocity of the wave being equal to the square root of the product of the depth into the acceleration of gravity, and the velocity acquired by a body falling freely through a height $d/2$ being $\sqrt{gh}$, it follows that the wave moves with the velocity a body would acquire in falling through a height equal to half the depth of the water.

279*. Velocity of a Wave of Compression or Dilatation in an Elastic Fluid.—We have next to consider the velocity with which a longitudinal wave (§ 266) moves. In such a wave the particles of the medium execute harmonic motions along straight lines which are parallel to the direction in which the wave is moving. By the forward motion of the particles the medium in front will be compressed, as shown in Fig. 216, so that owing to the elasticity of the medium the pressure will be increased, while by the backward motion the medium in front will be rarefied, and the pressure reduced.

Suppose that the medium is contained in a tube of which the cross section is unity, and that the velocity with which the wave moves is v, and that by imparting a velocity $-v$ to the medium, we bring all the waves to rest. Consider two imaginary partitions, A and B (Fig. 232), placed across the tube; then, since the waves are, by the motion of the

medium, rendered fixed in position, the phase of the wave at each of these partitions is always the same, and they always include the same number of waves between them. Thus we may suppose, if we like, that A is a crest or point where the compression is a maximum, and B a trough or point where the rarefaction is a maximum, and that there is a quarter wave between, so that we should be dealing with a similar case to that shown for water waves in Fig. 231.

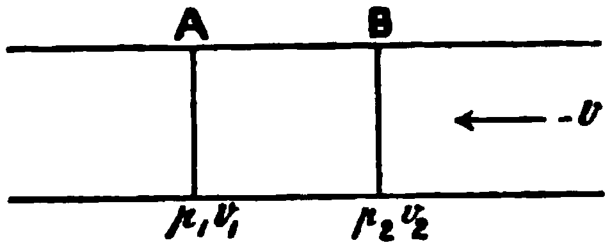


FIG. 232.

This being so, the pressure, and therefore also the volume of unit mass of the medium, and the velocity with which the particles of the medium are moving, due to their to-and-fro motion, remain constant both at A and at B. Let p_1, v_1, u_1 and p_2, v_2, u_2 be the pressure, volume of unit mass, and the velocity of the particles (supposing the medium were at rest) at A and B respectively.

Now since the space intercepted between A and B always contains the same number of waves, and the density of the medium at each point is constant, the mass of the medium intercepted between A and B is always the same. Hence the quantity, Q, of the medium which, owing to the motion supposed to be imparted to the medium and to the vibration of the particles, enters through B in a given time must be equal to that which leaves through A.

Now the velocity of the particles at A with reference to A is $u_1 - v$, so that the volume of the medium which passes through A in unit time is $(u_1 - v)$, for the cross section of the tube is unity. Since the volume of unit mass of the medium at A is v_1, the mass of the medium which crosses A in unit time is

$$Q=(u_1-v)/v_1 \quad . \quad . \quad . \quad (1).$$

In the same way the mass of the medium which enters through B in unit time is

$$Q=(u_2-v)/v_2 \quad . \quad . \quad . \quad (2).$$

Now the momentum lost through A due to this passage of the medium in unit time is the product of the mass lost by the velocity, or

$$Q(u_1 - v),$$

and that gained through B is

$$Q(u_2 - v).$$

Thus the *gain* of momentum within the space included between A and B, owing to the passage of the medium, is

$$Q(u_2 - u_1),$$

or, substituting for u_1 and u_2 the values given by (1) and (2), the gain is

$$Q^2(v_2 - v_1). \quad . \quad . \quad . \quad (3).$$

Now the momentum contained between A and B must remain constant throughout, for the state of the medium remains the same throughout. There must, therefore, be some cause which causes a loss of momentum exactly equal to the gain we have found above to be produced by the passage of the medium. This cause is the external forces, namely, the pressures at the two planes, which act on the medium contained between A and B. Now the pressure acting forwards on the medium contained within A and B, that is, the pressure at B, is p_2, while the pressure acting backwards on the front A is p_1. Hence the loss of momentum during unit time by the portion of the medium included between A and B, due to the external forces, is equal to the difference of the pressures,[1] that is,

$$p_1 - p_2.$$

If p_1 is greater than p_2, there will be a loss of momentum owing to the effect of the pressures, for the resultant force opposes the motion of the medium. But p_1 being greater than p_2, v_1 must be less than v_2, for at the greater pressure the volume of unit mass will be smaller. If v_1 is less than v_2, then $Q^2(v_2 - v_1)$ is positive, so that this quantity does really represent a *gain* of momentum, and we see that our signs are right.

Equating the gain and loss of momentum, we get

$$Q^2(v_2 - v_1) = p_1 - p_2,$$

or

$$Q^2 = \frac{p_1 - p_2}{v_2 - v_1}. \quad . \quad . \quad . \quad (4).$$

Now the elasticity of a substance is the ratio of the stress to the strain it produces (§ 122). Hence if E is the elasticity of the medium,

$$E = \frac{Stress}{Strain}.$$

Now we are considering the change in volume produced by a change in pressure. Hence if p_1 and p_2 are two pressures, and v_1 and v_2 the corresponding volumes, we have the stress is $p_1 - p_2$, and the strain is $(v_2 - v_1)/v_1$, for the strain is the change in volume per unit volume. Hence we get

$$v_1 E = \frac{p_1 - p_2}{v_2 - v_1}.$$

Therefore from (4)

$$Q^2 = v_1 E. \quad . \quad . \quad . \quad (5).$$

Now, in the case of all waves such as we are considering, the changes of pressure are very small compared to the whole pressure, so that v_1 is very nearly the same as v_0, the volume of unit mass of the medium when

1 The resultant force acting on a body is measured by the change in momentum it produces in unit time (Newton's Second Law, § 60). The force acting on the plane B is the pressure p_2 multiplied by the area, which is unity, while the force acting in the opposite direction on the plane A is in the same way p_1.

undisturbed by the passage of the wave. Also, if we consider a partition across the tube at the part of a wave where there is neither compression or rarefaction, that is, half-way between a crest and a trough, then, as is at once apparent from a consideration of the arrows in Fig. 216, the particles of the medium are at rest as far as their vibratory motion is concerned, and so the velocity of the medium at such a place is v, the velocity which has been impressed on the whole medium. Therefore

$$Q=v/v_0. \quad . \quad . \quad . \quad (6).$$

Hence from (5) and (6) $\qquad Q^2=v_0E=v^2/v_0^2,$

or $\qquad v^2=Ev_0,$

or $\qquad v=\sqrt{Ev_0}.$

Now v_0 is the volume of unit mass of the medium ; and if ρ is the density of the medium, we have

$$v_0=1/\rho.$$

Hence $\qquad v=\sqrt{E/\rho}.$

That is, the velocity of wave of compression and rarefaction in a medium is equal to the square root of the elasticity of the medium divided by its density. This is Newton's expression for the velocity of a longitudinal wave in a homogeneous medium.

Since the expression for the velocity does not involve the wave-length of the wave, it follows that waves of all wave-lengths or frequencies travel with the same velocity.

PART II—SOUND

CHAPTER II

PRODUCTION AND PROPAGATION OF SOUND

280. Sounding Body.—All bodies which are the source of that particular disturbance which, when it strikes our ear and affects the auditory nerves, we call sound, are in a state of vibration. This can be easily proved, for if a sounding body, such as a bell, is touched, the vibratory movement can be felt, and as under the influence of the resistance offered by the finger the vibrations die out, the sound emitted will also die out. The vibrations of a stretched string, which, when plucked, gives out a sound, are visible to the eye, and as they decrease in amplitude, the intensity of the sound also decreases.

281. Conveyance of Sound to the Ear.—In order that we may perceive a sound by our ears, it is necessary that the sounding body should be connected with our ear by an uninterrupted series of portions of elastic matter. The physical state of the matter, whether it is gaseous, liquid, or solid, is immaterial.

In order to show that this is the case, we may make the well-known experiment of suspending a bell by a thin string within the receiver of an air-pump. As the receiver is exhausted, the intensity of the sound heard when the bell is struck diminishes, till finally, when a fairly good vacuum is produced, no sound at all can be heard. If air, or even a few drops of a liquid which will form a vapour in the receiver, is introduced the sound is again heard, as is also the case if the sounding bell is allowed to dip into a vessel containing mercury or other liquid standing on the plate of the pump, or to touch a solid rod which is connected to the receiver or the plate. This experiment, therefore, shows that sound cannot be transmitted through a vacuum, but that it is transmitted through gases, vapours, liquids, and solids.

The ease with which sound travels through some solids, such as wood, is very clearly shown by holding one end of a long wooden rod against the ear, when even a very light scratch with a pin at the far end will be heard with great distinctness.

Since sound requires the presence of matter for its transmission, we are at once led to inquire what is the mechanism by which this transference takes place. There are two distinct methods by means of which

we might imagine that the sounding body affects the ear through an intervening space filled, say, with air. The sounding body might shoot out particles of some kind, and these particles, when they strike the ear, might cause the sensation of sound; or the sounding body may, on account of its vibratory motion, set up waves in the surrounding air, and these waves, travelling through the air, when they strike the ear may cause the sensation of sound. That the first explanation is untenable is shown by the experiment with the bell in the air-pump receiver, for the fact that no sound is heard when the air is removed shows that it is the air that conveys the sound; hence, if anything is projected out from the sounding bell, it must be air particles. We know, however, that glass is impervious to air particles, and hence, since we hear the bell through the glass, the sound cannot be conveyed to our ear by the projection of air particles from the sounding body. That the second method of transmission is the correct one will be abundantly evident as we proceed to describe the different phenomena, so that we need not insist on any special experiments by which it could be proved, but merely state that the sound is conveyed from the sounding body to the ear by waves which are set up by the sounding body, and which travel through the matter which connects the ear and the sounding body. Thus in the case of the bell in the receiver, when the latter is filled with air, the vibratory movements of the bell set up waves in the air, these waves travel out till they meet the glass walls of the receiver, in which, by their impact, they cause waves; these waves again travel through the glass, and at the outside cause waves in the surrounding air, and it is these latter which reach the ear.

As we have seen in § 266, there are two distinct kinds of waves which can be set up in a medium. The sound-waves are longitudinal waves in which the motion of each particle of the medium in which the wave is travelling moves backwards and forwards along a line in the direction in which the wave is travelling.

282. The Measurement of the Velocity of Sound in Air.—It is a matter of common experience that sound takes an appreciable time to travel from one place to another. Thus, when we observe a man breaking stones at a little distance, the sound of each blow is *heard* a very appreciable time after the blow is *seen* to be struck. In the same way, the puff of smoke and flash of a distant cannon is seen some time before the sound of the report is heard.

Since, as we shall see later on, the time taken by light to travel over a distance of a few miles is only a very small fraction of a second ($\frac{1}{186000}$ second for one mile), the most obvious way of measuring the velocity of sound in air is to note the interval which elapses between seeing the flash and hearing the report of a cannon. Then, if the distance of the cannon from the observer is known, the velocity of sound can immediately be calculated. The experiment is, however, not quite

as simple as it would at first sight appear. For the velocity of sound is affected by the wind, the temperature of the air, and the quantity of moisture in the air. In order to reduce the influence of these disturbing causes to a minimum, Aragó, and a number of French observers who worked with him, in 1822 chose two stations at a distance of 18612.52 metres apart. At each station a cannon was placed, and the observers were furnished with accurate chronometers. The cannon at the two stations were fired alternately, and the observers at the other station noted the time which elapsed between the flash and the report. In this way the effect of the wind on the velocity of sound was eliminated, for if the wind assisted the sound when travelling in one direction it would retard it when travelling in the opposite. As a result, they found that the sound took 54.4 seconds to travel in one direction and 54.8 seconds in the opposite. Hence the mean velocity was 340.9 metres per second. The temperature of the air was observed at a number of points between the two stations, the mean being $15^{\circ}.9$ C.

Experiments of this kind are, however, subject to a systematic error, due to the fact that it takes an appreciable time for an observer to become aware of a flash or a report, and to perform some action, such as pressing a button, to mark the time. If the same delay occurred in recording the flash and the report it would be of no consequence, since one would compensate for the other. This, however, is not the case, for not only does the time necessary to record a visual and an aural impression differ, but, what is more important, while the flash occurs unexpectedly, its occurrence acts as a warning to prepare for the report.

In order to eliminate this source of error, Regnault performed a series of experiments, in which the instant at which the discharge of a pistol took place was recorded on a rotating drum, by causing the bullet to cut a wire, and thus break an electric current. The arrival of the sound-wave was also automatically recorded by means of a stretched diaphragm, to the centre of which was attached a small metal disc. When the sound-wave struck this diaphragm it was forced back, and the metal disc touched a metal point, which had been adjusted so as very nearly to make contact with the disc. The disc and point formed part of a second electric circuit, and thus, when they came into contact, the circuit was completed, and a record was made on the rotating drum as before. The time which had elapsed between the two marks was determined by means of a series of marks made by the pendulum of a clock breaking another electric circuit, and by the trace made by a fine point attached to the prong of a tuning-fork. Regnault made observations of the velocity of sound in air by this method in long underground water-mains in Paris, in which no wind effects were to be feared, and where the temperature of the air could be more accurately determined than in the open. He found that the velocity depended to a slight extent on

the diameter of the pipe, and on the intensity of the sound. He also used the same apparatus for observations in the open air, and, as a result of all his experiments, came to the conclusion that the velocity of a sound of small intensity in dry air at a temperature of 0° C. was 330.6 metres per second.

In order to measure the velocity of sound in a limited distance, Bosscha, and after him Kœnig, used two small hammers, worked simultaneously by an electric current which was made by the pendulum of a clock, say every half-second. If the two instruments are at equal distances from the observer, the raps of the hammers will be heard simultaneously. If now, while one instrument is kept near the observer, the other is gradually moved away, the raps will no longer be *heard* simultaneously, for the sound from the further instrument will take longer to reach the ear than that from the nearer one. As the distance between the instruments is farther increased, a point will be reached such that the observer again hears both raps simultaneously; the further instrument being at such a distance that the sound of a rap from the nearer instrument reaches his ear at the same instant as the previous rap from the further one. Hence the sound has traversed the distance between the two instruments in the interval that elapses between two consecutive raps. If t is this interval of time, and d is the distance beween the instrument, then the velocity of sound is $\frac{d}{t}$.

The velocity of sounds of all pitches, that is, frequencies, was found by Regnault to be the same, for when a tune was played at one end of the long tube an observer at the other end heard the tune unaffected, except, of course, that the loudness was decreased. Now if, say, the high notes had travelled faster than the low, the interval of time which would elapse between the distant observer hearing a low note followed by a high would be greater than the proper interval, and *vice versa*. Thus the time of the tune, that is, the intervals of time between the successive notes, would be wrong at a distance from the place where it was played.

This experimental result agrees with the expression which we obtained in § 279 for the velocity of a longitudinal wave, for this expression does not involve the frequency of the wave, so that the velocity is independent of the frequency.

283. The Measurement of the Velocity of Sound in Water.— The velocity of sound in water was determined directly by Colladon and Sturm in the Lake of Geneva. Two boats were moored at a distance of 13,487 metres apart, and from one boat was suspended in the water a large bell. The hammer of the bell was worked by a lever, which was so arranged that at the instant it touched the bell a lighted match was caused to set fire to some gunpowder. The observer in the other boat was provided with a chronometer and a

large horn-shaped trumpet, the larger end of which was closed by an india-rubber membrane and dipped into the water. He noted the time which elapsed between the flash of the powder and the sound of the bell which reached him through the water. In this way they found that the velocity of sound in water was 1435 metres per second at a temperature of 8°.1 C.

284. The Measurement of the Velocity of Sound in Solids.—A direct determination of the ratio of the velocity of sound in air to that in cast iron was made by Biot, by measuring the interval between the two sounds heard when the end of a long cast-iron pipe was struck by a hammer. One of these sounds, the first, is that which travels through the iron, and the other that which has travelled through the air in the pipe. In this way he found that the velocity in cast iron was 10.5 times that in air, but the experiments were not of any high degree of accuracy, since the interval between the two sounds (2.5 seconds) was very small, and therefore hard to measure accurately, and the pipe consisted of a number of separate pieces joined together by lead, so that it was not a continuous rod of iron.

We shall see in a subsequent section how the velocity of sound in solids can be indirectly determined with a much higher degree of accuracy.

285. Calculation of the Velocity of Sound in a Homogeneous Medium.—In the case of longitudinal waves, such as occur in the case of sound, Newton first showed, as we have done in § 279, that the velocity v is given by the equation

$$v=\sqrt{\frac{E}{\rho}},$$

where E is the elasticity, and ρ the density of the medium.

Further, we have seen in § 122 that the elasticity E is equal to the quotient $\frac{stress}{strain}$.

In the case of a gas the stress is the excess, or the defect, of the pressure at a given point and at a given time during the passage of the wave, over the average pressure at that point, or, what comes to the same thing, over the pressure when the wave is not passing.

If P is the undisturbed pressure, or mean pressure, p the increase of pressure, and V and v the original and change in volume of unit mass respectively, then the stress is p and the strain is v/V, since v is the total deformation and v/V is the deformation per unit volume (§ 130). Hence

$$E=\frac{p}{v/V}=\frac{pV}{v}.$$

Now if Boyle's law holds with regard to the compressions and rare-

factions which take place when the sound-waves travel through a gas, we have

$$PV=(P+p)(V-v),$$
$$=PV-Pv+pV-vp.$$

Since in the case of sound-waves the change in pressure p, and therefore also the change in volume v, is excessively small, the product pv will be very small indeed. Hence, neglecting this small quantity, we get

$$Pv=pV,$$

or $$P=\frac{pV}{v}.$$

But $$E=\frac{pV}{v}.$$

Hence $$E=P.$$

Thus if the changes of pressure and volume in the gas obey Boyle's law, which of course will only hold if the temperature remains constant during the compression and rarefaction, the elasticity is numerically equal to the pressure. Substituting in Newton's equation for the velocity, this becomes

$$v=\sqrt{\frac{P}{\rho}}.$$

In the case of air at a pressure of one atmosphere (1013300 dynes per square cm.) and at a temperature of 0° C., the density ρ is 0.001293. Hence, substituting these values in the equation for v, we have

$$v=\sqrt{\frac{1013300}{0{\cdot}001293}}=28026 \text{ cm. per second.}$$

Now we have seen that the velocity of sound in air, as found by experiment under the above condition of pressure and temperature, is 33060 cm. per second. The difference between the observed and calculated values being much greater than any possible error of experiment, we are led to the conclusion that some of the assumptions made above are erroneous.

Now we have seen in § 258 that when a gas is compressed its temperature rises, and when it is allowed to expand the temperature falls. In the above reasoning we have supposed that the compressions and rarefactions that take place when a sound-wave traverses a gas are so slow that, by conduction from and to the surrounding air, the temperature of the compressed or rarefied air remains constant. If, however, the compressions and rarefactions take place with such rapidity that the air has not time to lose heat when it is warmed by compression, or to gain heat from surrounding particles when it is cooled by the expansion

during the passage of a wave, then the compression and expansion will be adiabatic (§ 258).

Now in the case of an adiabatic compression or expansion the volume and pressure are connected by the relation

$$PV^k = \text{constant},$$

where k is the ratio of the specific heat at constant pressure to the specific heat at constant volume.

Hence, making the assumption that the changes of pressure and volume caused by a sound-wave are adiabatic, and not isothermal, as we previously assumed, we have, using the same notation as before,

$$PV^k = (P+p)(V-v)^k.$$

Expanding $(V-v)^k$ by the binomial theorem, and, since v is small, neglecting all terms which involve v^2 or higher powers of v, we get

$$\begin{aligned} PV &= (P+p)(V^k - kV^{k-1}v) \\ &= VP^k - k.PV^{k-1}v + pV^k - kV^{k-1}vp, \end{aligned}$$

or, neglecting the term involving the product of the small quantities p and v,

$$kPV^{k-1}v = pV^k,$$

or, dividing both sides by V^{k-1},

$$kPv = pV$$

$$\therefore \qquad kP = \frac{pV}{v},$$

or since, $\frac{pV}{v} = E$, we have

$$E = kP,$$

and hence the equation for the velocity of sound becomes

$$v = \sqrt{\frac{Pk}{\rho}}.$$

In the case of air $k = 1.41$, and hence the calculated value of the velocity is $28026 \times \sqrt{1.41} = 33240$ cm. per second, which agrees fairly well with the observed value.

The above is Laplace's formula for the velocity of sound in a gas, and the fact that the calculated value of the velocity agrees with the observed value is a proof of the accuracy of the assumptions on which it is based. We shall, indeed, see further on that the most accurate method of determining the ratio of the specific heats for a gas is to measure the velocity of sound in the gas, and to calculate the value of k from Laplace's equation for the velocity.

In the case of liquids, the compressibility is so small that the thermal changes which take place on this account have no appreciable effect on the elasticity, so that in this case Newton's equation is applicable. For water we have at 4° C. $\rho = 1$, and an increase of pressure of one atmosphere causes unit volume to decrease by .000049 units of volume. Hence in the *c.g.s.* system the elasticity is given by

$$\frac{stress}{strain} = \frac{760 \times 981 \times 13.59}{0.000049},$$

and

$$v = \sqrt{\frac{760 \times 981 \times 13.59}{0.000049}}$$

$$= 142500 \text{ cm. per second,}$$

a result which agrees fairly well with the observed value.

286. Effect of Temperature on the Velocity of Sound.—We have seen in § 197 that in a case of a gas $\frac{PV}{1+at}$ is a constant. Hence if P_0 is the standard pressure, and V_0 is the volume of unit mass of a gas at this pressure and at a temperature of 0°, we have

$$\frac{PV}{1+at} = P_0 V_0.$$

But if V is the volume of unit mass, we have, since the density ρ is the mass of unit volume, the relation $V = 1/\rho$; and hence

$$\frac{P}{\rho(1+at)} = \frac{P_0}{\rho_0},$$

or

$$\frac{P}{\rho} = \frac{P_0}{\rho_0}(1+at).$$

Hence, if v_t is the velocity of sound at a temperature t, we have

$$v_t = \sqrt{k\frac{P}{\rho}}$$

$$= \sqrt{k\frac{P_0}{\rho_0}(1+at)},$$

$$= v_0\sqrt{(1+at)}.$$

For if v_0 is the velocity of sound at 0°, and under standard pressure,

$$v_0 = \sqrt{k\frac{P_0}{\rho_0}}.$$

Now if we expand $\sqrt{(1+at)}$ by the binomial theorem, and, since a is small, neglect terms in a^2 and in higher powers of a, we get

$$v_t = v_0\left(1+\frac{at}{2}\right),$$

or, since $a = .00366$,

$$v_t = v_0(1+.00183t).$$

In the case of air $v_0 = 33060$ cm./sec.
Hence

$$v_t = 33060 + 60.5t \text{ cm./sec.}$$

Changes of pressure unaccompanied by changes of temperature, such as a change in the barometric pressure, will not affect the velocity of sound in a gas, for by Boyle's law

$$PV = P_0V_0$$

or

$$\frac{P}{\rho} = \frac{P_0}{\rho_0}.$$

Hence the change in pressure effects the elasticity and the density in the same ratio, so that the velocity of sound is unaffected.

CHAPTER III

PITCH—MUSICAL SCALE

287. Quality of Sounds.—Sounds which affect our ear are divided into two classes. One of these consists of short sounds which last only for a short time, or, if they last for any length of time, are continually changing their character, and are called noises. The other class consists of sounds the character of which is that the vibrations by which they are caused are periodic; these are called musical notes.

Musical notes differ from one another in three important particulars: (1) They may be of different intensity or loudness. Thus when we move away from a sounding body the intensity of the note given by the body decreases, but does not otherwise alter. (2) The pitch of two notes may be different. We shall see that the pitch or highness of a note depends on the frequency of the vibrations of the sounding body. (3) The notes given out by two different instruments, such as the cornet and the piano, although they may be of the same pitch and intensity, are clearly distinguishable by the ear. This quality of a musical note is called its *timbre*.

288. Pitch of a Note.—That the pitch of a sound depends on the frequency, or the number of vibrations per second made by the sounding body, can be shown by the instrument called a syren. The usual form of this instrument is shown in Fig. 233. It consists of a circular disc BB′, mounted on a vertical spindle D, so that it turns freely. This disc is pierced by a number of holes at equal distances apart. The disc is pivoted so that it just clears the upper surface AA′ of a small wind-box, H, which is connected with a bellows, by means of which a continuous current of air can be forced into the instrument. The plate A is pierced by the same number of holes as the movable disc. The holes in the fixed and movable discs are not drilled at right angles to the surface of the plates, but those in A and B are inclined in opposite directions, as shown at *a* and *b*. Hence

FIG. 233.

the air, when forced out through a, strikes against the side of the hole b, and causes the disc to rotate in the direction of the arrow. Each time the holes in the upper disc come opposite the holes in the lower plate, a puff of air escapes, and the disc receives an impulse. If the disc is rotating sufficiently rapidly, these puffs will produce a musical note, and as the velocity of rotation of the disc, and therefore also the frequency of the puffs increases, the pitch of the note rises.

The syren also permits of the determination of the frequency of a musical note, for if there are x holes in the upper and lower plates, then, during a complete revolution of the upper plate, a hole in the upper plate will coincide with a hole in the lower plate x times. For the angular distance between two holes in the lower plate is $360/x$, and hence, when the upper plate has turned through this angle, each hole on the upper plate will have just moved on one, and will coincide with the next hole in the lower plate. The number of coincidences during one turn, or 360°, is therefore

$$360° \div 360/x \text{ or } x.$$

If the movable plate makes n revolutions per second, the number of puffs per second, or the frequency of the sound, will be nx.

In performing the experiment the pressure of the wind is increased, thus causing the movable plate to rotate faster and faster, till the pitch of the note emitted is the same as that of the note whose frequency has to be measured. The wind pressure is then kept constant, and the number of revolutions made by the movable plate in a given time is obtained by means of the toothed wheels R and S, which can be put into gear with the endless screw, V, attached to the spindle by pressing on the knob E, at the commencement of the interval, and put out of gear at the end of the interval by pressing on the knob F. The wheel R moves on one tooth for each revolution of the disc, and has 100 teeth. The wheel S is moved on one tooth every time R completes a revolution. Hence the number of turns and hundreds of turns can be read off on two dials by means of pointers attached to R and S.

In performing such an experiment, there is considerable difficulty in keeping the speed of rotation constant, and such that the note given by the syren is in unison with the note whose frequency is being measured. For this reason, the more modern forms of syren are driven by a small electric motor, and not by the pressure of the escaping air in the inclined holes. The speed of the motor is kept constant by means of an electric regulator.

Other methods of measuring the pitch of the musical note given out by a sounding body will be considered in subsequent sections.

289. The Musical Scale.—We have in the previous section referred to a note as being higher or lower than another, and this statement has probably conveyed the required impression to all readers. We

have now to consider the physical connection between sounds of various pitch in their relation to the pleasing, or otherwise, effect they produce on our ear.

It is found that, whether we are dealing with the consecutive sounding of two notes (melody), or with the simultaneous sounding of two notes (harmony), the ear only takes cognisance as forming pleasing combinations of notes in which the *ratio* of the frequencies is expressible by two integers which are neither of them very large. Hence, in considering the relations between the pitches of musical notes, we have to deal with the ratio of their frequencies, and not with the difference. The ratio of the frequencies of two notes is called the *interval* between the notes.

Hence if the frequencies of three notes are n_1, n_2, and n_3, the interval between the first and second is n_1/n_2, and the interval between the second and third is n_2/n_3. The interval between the first and third is n_1/n_3, and since $n_1/n_3 = n_1/n_2 \times n_2/n_3$, we see that the interval between the first and third is obtained by *multiplying* the interval between the first and second by that between the second and third. Thus the "sum" of two intervals is obtained by multiplying the intervals together, for we may look upon the interval n_1/n_3 as being made up of the two separate intervals n_1/n_2 and n_2/n_3, which will take us from the note n_1 to the note n_3. In the same way, since $n_1/n_3 \div n_1/n_2 = n_2/n_3$, the difference between the intervals n_1/n_3 and n_1/n_2 is obtained by dividing one of these intervals by the other. Thus the interval n_1/n_3 is greater than the interval n_1/n_2 by the interval n_2/n_3.

We have now to consider the intervals between notes which convey to the ear certain well-known and distinctive sensations, independent of the absolute frequency of the two notes. In the first place, it is found that all notes of which the interval is 1, *i.e.* all notes having the same frequency, although they may have very different intensities and timbres, are yet clearly recognised, whatever their absolute frequency may be, as having some quality in common, that is, they all have the same pitch.

Next, if the interval between two notes is $\frac{1}{2}$, that is, if the frequency of one note is twice that of the other, the two notes when sounded, either consecutively or together, produce a not unpleasing sensation on the ear, that is, they are said to be in accord or in consonance. This interval is called an *octave*. In this case again, as in fact in all cases, the ear recognises this relation between two notes whatever be the absolute frequency of the notes. Thus two notes of which the frequencies are 256 and 512 form an interval of an octave, as also do notes having the frequencies 128 and 256, 370 and 740, &c.

Between a given note and its octave the ear recognises a definite succession of notes of which the frequencies are well defined. These notes form what is called the *musical scale*. Starting from a note of *any*

frequency, we can construct a scale, and the interval between successive notes will in all cases be the same. Including the lowest note, which is called the tonic, and its octave, the scale consists of eight notes. The notes of the scale are generally indicated by the letters C, D, E, F, G, A, B, *c*, or the names do, re, mi, fa, so, la, si, do_1. The interval between the tonic and each note in the scale, as well as the interval between each two consecutive notes, is shown in the following table :—

C		D		E		F		G		A		B		*c*
do		re		mi		fa		so		la		si		do_1
1		$\frac{9}{8}$		$\frac{5}{4}$		$\frac{4}{3}$		$\frac{3}{2}$		$\frac{5}{3}$		$\frac{15}{8}$		2
	$\frac{9}{8}$		$\frac{10}{9}$		$\frac{16}{15}$		$\frac{9}{8}$		$\frac{10}{9}$		$\frac{9}{8}$		$\frac{16}{15}$	

For the sake of remembering the relative frequency of the notes of the scale, the following series of whole numbers, which are proportional to the frequencies, is useful :—

C	D	E	F	G	A	B	*c*
24	27	30	32	36	40	45	48

It will be seen that, considering the intervals between consecutive notes of the scale, there are three separate and distinct intervals, and these intervals have received special names. The interval $\frac{9}{8}$ or 1.125 is called a *major tone*, the interval $\frac{10}{9}$ or 1.111 is called a *minor tone*, and the interval $\frac{16}{15}$ or 1.067 is called a *limma*.

The difference between a major tone and a minor tone is $\frac{9}{8} \div \frac{10}{9}$ or $\frac{81}{80}$, and is called a *comma;* while the difference between a minor tone and a limma is $\frac{10}{9} \div \frac{16}{15}$ or $\frac{25}{24}$, and is called a *diesis*.

If any two notes of the scale are sounded together, the ear recognises that the combinations are some of them more consonant than the others. The most consonant interval is the octave or $\frac{2}{1}$; next to this comes the interval between the tonic and the fifth note of the scale, namely G, for which the interval is $\frac{3}{2}$. This interval is called a *fifth*, since it occurs between the first and fifth notes of the scale. The next most complete consonance is between the tonic and the fourth note, or F, for which the interval is $\frac{4}{3}$, and this interval is called a *fourth*. Then we have in succession, as far as consonance is concerned, the following :—

Major-sixth interval, $\frac{5}{3}$, between C and A.
Major-third ,, $\frac{5}{4}$, ,, C ,, E.
Minor-third ,, $\frac{6}{5}$, ,, E ,, G.

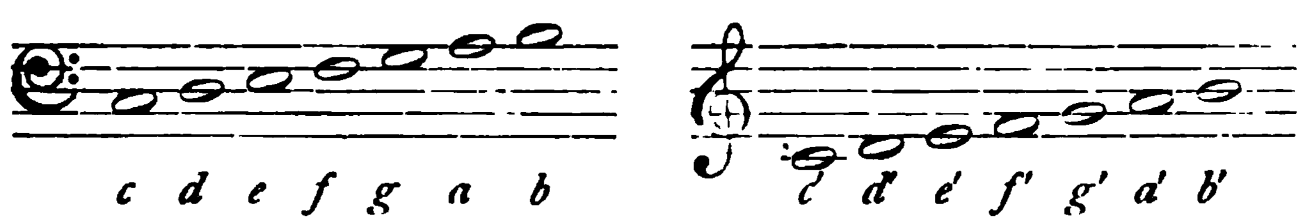
In addition to the interval between the tonic and the fifth note above, namely G, which is sometimes called the *dominant*, there are other intervals between a note and the fifth one above it. These intervals do not differ from $\frac{3}{2}$ by more than $\frac{81}{80}$, or a comma, except that between B

and f (in the octave above), which exceeds $\frac{3}{2}$ by the interval $\frac{25}{24}$, or a diesis, and is called a minor fifth. Hence with this exception, since the interval of a comma can barely be appreciated by the ear, all these intervals are very consonant. In the same way, with one exception, the fourths are all the same to within a comma. The thirds and sixths are either major, in which the interval does not differ by more than a comma from $\frac{5}{4}$ or $\frac{5}{3}$, as the case may be, or minor, in which the interval is less by a diesis $\frac{25}{24}$.

In addition to the notes given above, use is made in music of additional notes which are derived from the above by either raising or lowering the pitch of each note by a diesis, *i.e.* $\frac{25}{24}$. If the pitch is raised by a diesis the note is said to be sharpened. Thus if a note, say G, of which the frequency is 384 is increased to $384 \times \frac{25}{24}$, or 400, the new note is called G sharp, and is indicated by the symbol G♯. In the same way the note having the frequency $384 \times \frac{24}{25}$, or 368.6, is called G flat, and is indicated by G♭.

In order to distinguish the notes of the scale in the various octaves, it is usual to use accented letters. The lowest octave, namely, that in which the C makes 33 vibrations per second, is indicated by the letters C_1 to B_1. This octave is often called the 16-foot octave, since an open organ-pipe 16 feet long gives a note included in this octave. The next higher octave is indicated by the unaccented capital letters C to B, and is called the 8-foot octave. The next octave is indicated by the unaccented small letters c to b, and is called the 4-foot octave. The remaining octaves are indicated by the letters c' to b', c'' to b'', c''' to b'''.

The relation of the above octaves to the ordinary musical notation is shown below :—



We have hitherto said nothing as to the absolute frequency of any of the notes of the scale, for the intervals between the notes are quite independent of the absolute frequencies, and only depend on the ratios of these quantities. The standard pitch adopted in different countries varies considerably. Thus the French standard pitch is 435 for the a', the German is 400, and the concert-pitch, as it is called, is 460.

For many purposes, particularly in physics, it is convenient to take as the standard 426.66 for a', since then the frequency of c' will be 256, which number is a power of two. Taking the frequency of c' as 256, the frequencies of the other notes of the scale will be as follows :—

c'	d'	e'	f'	g'	a'	b'	c''
256	288	320	341.3	384	426.7	480	512

As it is often convenient to be able immediately to obtain the frequency of a note in any octave, the following table is given, in which c' is taken as having a frequency of 264, this being the standard adopted by the Stuttgart Congress and the Society of Arts :—

Notes.	Octave.						
	C_1 to B_1.	C to B.	c to b.	c' to b'.	c'' to b''.	c''' to b'''.	c'''' to b''''.
C, do . .	33	66	132	264	528	1056	2112
D, re . .	37.125	74.25	148.5	297	594	1188	2376
E, mi . .	41.25	82.5	165	330	660	1320	2640
F, fa . . .	44	88	176	352	704	1408	2816
G, sol . .	49.5	99	198	396	792	1584	3168
A, la . . .	55	110	220	440	880	1760	3520
B, si . . .	61.875	123.75	247.5	495	990	1980	3960

Since the interval between two notes of which the frequencies are m and n is given by the ratio m/n, if we take the logarithms of the frequencies, the difference of the logarithms will be the logarithm of the interval. For

$$\log (m/n) = \log m - \log n.$$

Hence every interval has the same logarithm, no matter what the absolute pitches of the two notes, so that if we require to determine the frequency of a note which will form a given interval with a given note, all we have to do is to add the logarithm of the interval to the logarithm of the frequency of the given note, and the sum will be the logarithm of the frequency of the required note. Thus from the table below we see that the logarithms corresponding to the interval of a fifth is 0.17609. If then we wish to find the frequency of a note making a fifth with one of 256 vibrations, we have, if x is the frequency required,

$$\frac{x}{256} = \frac{3}{2}.$$

Or, taking logarithms,

$$\begin{aligned} \log x &= (\log 3 - \log 2) + \log 256 \\ &= 0.17609 + 0.40824 \\ &= 0.58433 = \log (384). \end{aligned}$$

This example will make the reason for the rule clear.

The logarithms of the various intervals are given in the following table, as well as the note which, together with the tonic C, will give the interval :—

Notes.	Intervals.		Logarithm. Natural Scale.	Logarithm. Equally Tempered Scale.
C	Unison	$\frac{1}{1}$	0.00000	0.00000
	Comma	$\frac{81}{80}$	0.00540	
C♯	Semitone or diesis	$\frac{25}{24}$	0.01773	
	Limma	$\frac{16}{15}$	0.02803	0.02509
D♭	Minor second	$\frac{27}{25}$	0.03342	
	Minor tone	$\frac{10}{9}$	0.04576	
D	Major second or major tone	$\frac{9}{8}$	0.05115	0.05017
D♯	Augmented second	$\frac{75}{64}$	0.06888	
E♭	Minor third	$\frac{6}{5}$	0.07918	0.07526
E	Major third	$\frac{5}{4}$	0.09691	
F♭	Minor fourth	$\frac{32}{25}$	0.10721	0.10034
E♯	Augmented third	$\frac{125}{96}$	0.11464	
F	Perfect fourth	$\frac{4}{3}$	0.12494	0.12543
F♯	Augmented fourth	$\frac{25}{18}$	0.14267	
G♭	Minor fifth	$\frac{36}{25}$	0.15836	0.15052
G	Perfect fifth	$\frac{3}{2}$	0.17609	0.17560
G♯	Augmented fifth	$\frac{25}{16}$	0.19382	
A♭	Minor sixth	$\frac{8}{5}$	0.20412	0.20069
A	Major sixth	$\frac{5}{3}$	0.22185	0.22577
A♯	Augmented sixth	$\frac{125}{72}$	0.23958	
B♭	Minor seventh	$\frac{9}{5}$	0.25527	0.25086
B	Major seventh	$\frac{15}{8}$	0.27300	
c♭	Minor octave	$\frac{48}{25}$	0.28330	0.27595
B♯	Augmented seventh	$\frac{125}{64}$	0.29073	
c	Octave	$\frac{2}{1}$	0.30103	0.30103

290. Temperament.—In music it is often necessary to use scales having different tonics. Suppose then we require to form a scale of which the tonic has a frequency of 330, *i.e.* corresponds to the e' of the scale. We are at once met with the difficulty that the notes of the old scale will not fit into the new scale. For instance, the interval between e' and f' in the old scale is $\frac{16}{15}$, while the interval between the tonic and the next note above ought to be $\frac{9}{8}$.

The same difficulty remains even if we deal with the more extended scale in which each note is sharpened and flattened, for, as is shown by the table of the logarithms of the intervals given above, the sharp of one note does not necessarily have the same frequency as the flat of the note above. We thus see that the notes belonging to any given key will not serve as consecutive notes in any other key. Of course this would, in the case of an instrument such as the piano, in which the pitch of the various notes is fixed, render it practically impossible to arrange for more than one key. Hence, in order to be able to use the same series of notes for music written in different keys, the relative frequencies of the various

notes are in practice so altered that the notes belonging to the scale in any one key may be used for the scale in any other key, without any of them differing from the true scale by more than do the notes of the original scale. This process of adjusting the notes of the scale is called *temperament*.

There are two methods of temperament in common use. In the one the more consonant intervals, such as the fifth, are kept accurate, and the errors due to temperament are spread over the remaining intervals; this method is called unequal temperament. In the other method the interval of the octave is kept correct, and the errors are spread equally over the remaining intervals; this is called equal temperament, and is that usually adopted in the piano.

The interval of a major tone is equal to $\frac{9}{8}$ or 1.125; that of a minor tone is equal to $\frac{10}{9}$ or 1.111; and that of a limma is equal to $\frac{16}{15}$ or 1.067. Hence a major and a minor tone are very nearly the same, while each of these is very nearly equal to the interval of two limmas, since two limmas are equal to $(\frac{16}{15})^2$ or 1.138.

In order to obtain an equally tempered scale, the above approximate relations are assumed to be exactly true, that is, we take the major tone, the minor tone, and the two limmas as being the same.

If the interval of a tempered limma is called x, then there will be twelve of these limmas in the octave, and since an interval of twelve limmas is equal to x^{12}, for the interval between two notes is equal to the product of the intervals between the intermediate notes, while the interval of an octave is 2, we shall have in the equally tempered scale

$$x^{12}=2,$$

or

$$x=2^{\frac{1}{12}}=1.059.$$

In this tempered scale the limma is 1.059, instead of being 1.067, as it is in the true scale, while the tempered major and minor tones are each equal to 1.122, instead of being 1.125 and 1.111 respectively.

In the following table, the relative frequencies of the notes on the natural and the equally tempered scales are shown :—

	C	D	E	F	G	A	B	c
Natural Scale . .	1.000	1.125	1.250	1.333	1.500	1.667	1.875	2.000
Tempered Scale .	1.000	1.122	1.260	1.325	1.498	1 682	1.888	2.000

On the equally tempered scale C♯ and D♭ are the same notes, and this relation also holds with regard to D♯ and E♭, F♯ and G♭, G♯ and A♭, A♯ and B♭, while E♯ is the same as F, and F♭ is the same as E, as is also the case for B♯ and C♭. This is at once evident, for the interval from C to D is on the tempered scale 1.122 or $(1.059)^2$, while the interval between C and C♯ is 1.059, and that between D♭ and D is also 1.059. These relations are also clearly evident from the last column of

the table on page 381, where brackets show the various intervals which are taken as equal on the equally tempered scale.

291. Tones — Harmonics — Overtones.—When a single note is sounded on many kinds of musical instruments, a practised ear can detect that, in addition to the note the frequency of which corresponds to the note sounded, there are present notes corresponding to other frequencies, though these are very much less intense than the principal note. A note which the ear cannot break up in this manner is called a *tone*. Thus musical notes are in general composed of tones, the pitch of the note being that corresponding to the lowest tone it contains.

If n is the frequency of a tone, then the tones of which the frequencies are $2n$, $3n$, $4n$, &c., are called the *harmonics* of the tone n, and this tone is called the *fundamental*.

The tones which go to build up a note are not necessarily the harmonics of the lowest tone, so, for distinction, they are called the *overtones* or *upper partials* of the fundamental.

In the case of tones the vibrations of the sounding body, as well as the waves produced in the air, are simple harmonic vibrations ; and it is from this fact, first discovered by Ohm, that the name harmonic vibration is derived.

CHAPTER IV

REFLECTION, REFRACTION, AND INTERFERENCE

292. Reflection of Sound.—We have a familiar case of the reflection of sound in the echo, which is due to the reflection of the sound-waves by some large vertical surface, such as a cliff or the side of a house.

The reflection of sound may also be shown by means of what is called a sensitive flame, which consists of a gas flame produced by a pin-hole burner, in which the pressure of the gas has been increased till the flame is on the point of flaring. Such a flame forms a very sensitive detector of sound-waves, particularly those of a very high pitch. The form of the flame when unaffected is shown at B (Fig. 234), while, when a sound of suitable pitch is produced, the flame flares and shortens into the shape shown at A. The sensitive flame is placed at A (Fig. 235), in front of a tube CD, and a whistle B is placed opposite the end of another tube EF, while a screen is placed at GH, so as to screen off the direct action of the whistle on the flame. The pressure of the gas is adjusted till the flame just does not flare when no reflector is placed at I; it will then be found that, on placing a reflecting surface at I in such a way as to be equally inclined to the axes of the two tubes, the flame immediately begins to flare, and continues to do so as long as the reflecting surface remains in position.

FIG. 234.

The direction in which the sound-wave proceeds after reflection can be obtained in exactly the same way as that adopted in § 274; and, as there, it will be found that the angle of incidence is always equal to the angle of reflection. This fact is also proved by the experiment described above, for it is only when the normal to the

reflector I bisects the angle between the axes of the tubes that the flame flares ; and when the normal bisects this angle, the angle of reflection is equal to the angle of incidence.

Use is made of the reflection of sound in several well-known instruments. Thus in the ear-trumpet, the sound-waves that are caught by the bell-shaped mouth are reflected from the sides of the trumpet, and the cross section of the wave-front is decreased up to the ear-end. Thus the amplitude of the sound-waves increases as the cross section of the air in which they take place is decreased, and the amplitude of the waves when they strike the ear is much greater than if this concentration did not take place.

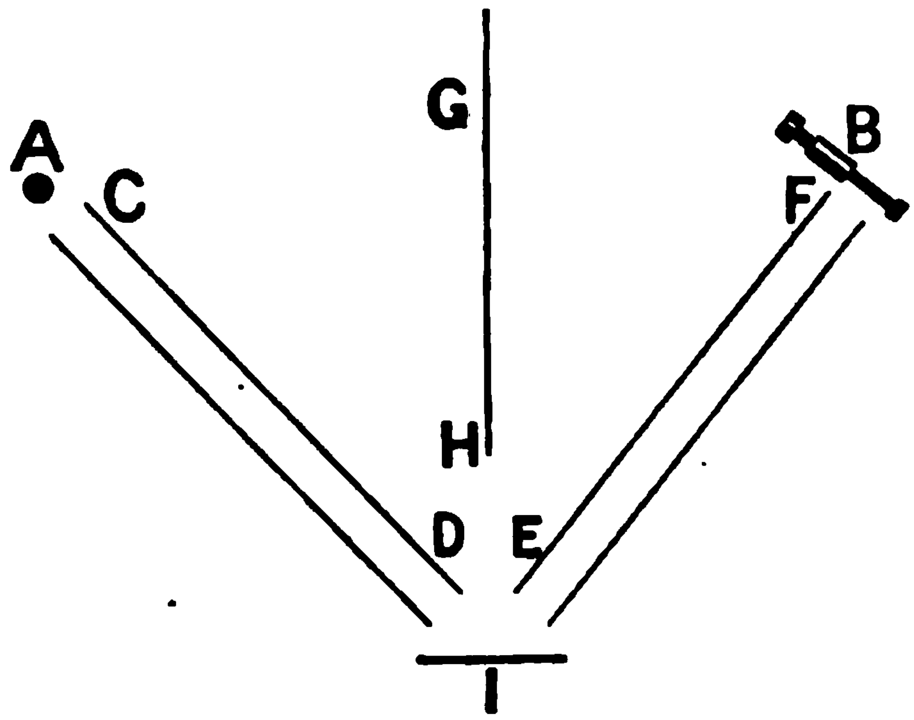


FIG. 235.

The sounding - board placed over the head of the speaker in large halls is another application of the reflection of sound. It consists of a reflecting surface placed so as to reflect those sound-waves that strike it down towards the audience. Hence the waves, that would otherwise spread up to the roof and be irregularly reflected there, are directed downwards, and assist in making the speaker audible.

In the case of speaking-tubes, the sound-waves, instead of spreading out in spheres, as they would do in the open air, are, by reflection at the sides of the tube, confined within the tube, so that they travel forward with comparatively small decrease in amplitude, the wave-front remaining of the same cross section throughout. A similar effect is produced when a watch is held against one end of a long wooden rod, and the ear is held against the other end. The ticking of the watch can be heard almost as clearly as if it were held close to the ear. The reason is that the sound-waves in the wood, when they reach the bounding surface between wood and air, are almost entirely reflected, and thus the wave proceeds down the rod without much of the energy escaping into the surrounding air.

The difficulty of hearing sounds at a distance on certain days is supposed to be due to the fact that on such acoustically opaque days there exist columns and layers of air at different temperatures, and that the sound-waves get partly reflected at each passage from air at one temperature to air at another. Such reflection must occur, for the velocity with which the sound-waves travel will be different if the temperature of

the air is different, and whenever a wave passes from one medium to another, in which it moves with a different velocity, it is partly reflected at the boundary between the two media.

293. Refraction of Sound.—When a sound-wave passes from one medium to another, the direction in which the sound-wave is travelling is in general altered, and is said to be refracted. The laws which govern the refraction of sound are the same as those in the case of light (see § 341). In the case of sound, however, it is difficult to obtain quantitative results. The fact that sound can be refracted may, however, be shown by having a lens-shaped india-rubber bag filled with a gas other than air. If the lens is convex (§ 348) and is filled with carbon dioxide, a gas in which sound travels with a smaller velocity than in air, the sound-waves will be brought to a focus, so that by placing a whistle at one side and a sensitive flame at the other, it will be possible to arrange matters so that the flame does not flare when the lens is not interposed, but does when the lens focuses the waves on the flame. If the lens is filled with hydrogen, a gas in which sound travels more rapidly than in air, then the convex lens will act as a diverging lens (§ 348), and a somewhat similar experiment can be performed in which the flame flares without the lens, but, owing to the spreading of the sound-waves by the lens, does not do so when the lens is interposed.

Sound-waves are often refracted on account of the motion of the wind. Thus suppose we have a plane sound-wave, in which the wave-front is a vertical plane, moving against the wind. Near the surface of the earth the velocity of the wind is in general less than at some distance above the surface. Now the sound-wave will travel at the same speed through the air, but, owing to the contrary motion of the air, the distance moved through by the wave-front relative to the earth will be greater near the surface of the earth than higher up. The wave-front will therefore become inclined, the top lagging behind the bottom, and, since the motion of the wave is at right angles to the wave-front, the direction of motion of the wave, instead of being parallel to the surface, will be inclined upwards. Thus the sound-waves may pass over the head of an observer who is on the windward side of the place where the sounds originate. When the sound is travelling with the wind the opposite effect is produced, the waves being refracted downwards. This effect partly accounts for the greater distance sounds can be heard when the sound is moving with the wind, than when the sound is moving against the wind.

Another cause of the refraction of sound-waves is the unequal heating of the various strata of air. In general during the day the air near the ground is hottest. As sound travels quicker in hot than in cold air, the result is that the waves get refracted upwards.

294. Interference of Sound-Waves.—The fact that sound-waves can interfere may be easily shown, and the wave-length of the note used

measured (only roughly it is true) by means of the instrument shown in section in Fig. 236. A whistle or reed is sounded at A, and the sound-waves can reach the sensitive flame placed at F by the two tubes ABC and ADC. If these two paths are of equal length, then the waves that reach C after travelling by the two tubes will be in the same phase, and hence they will strengthen one another. One tube, D, can be lengthened, and hence the path which the sound has to traverse can be made longer in this tube than the other. If the tube is gradually pulled out, a position such as D′ can be obtained, for which the waves that travel by the two paths are in opposite phase when they reach C, and hence they interfere, so that the flame is unaffected although the whistle is sounding as strongly as before. When this occurs the path AD′C traversed by one wave is longer than the path ABC traversed by the other by half a wave-length, so that when a crest reaches C *via* B, the preceding trough which has travelled *via* D′ has only just reached C, and these two neutralise one another, so that the air at C is undisturbed.

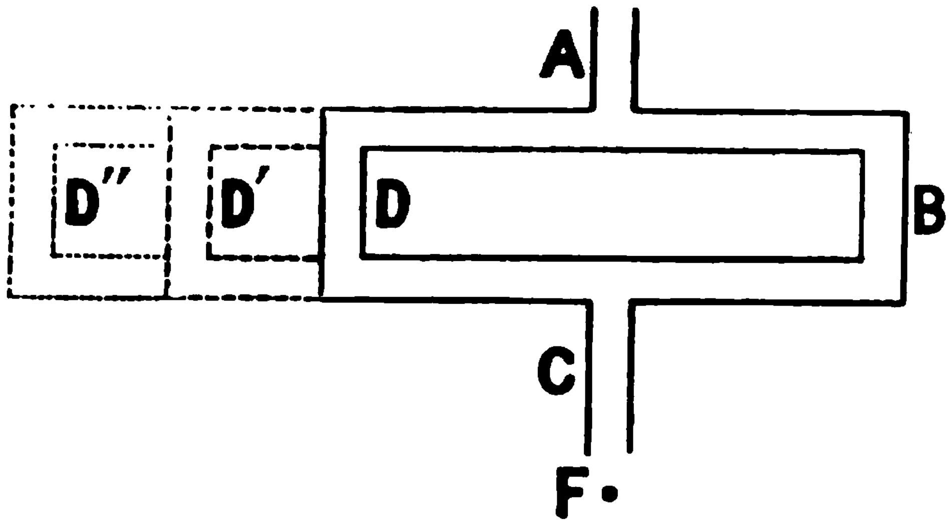


FIG. 236.

If the tube D is pulled further out, the paths differ by more than half a wave-length, and hence the two sets of waves commence to strengthen one another again, till when the tube is pulled out to D″, and the difference between the paths amounts to a whole wave-length, the flame is almost as strongly affected as it was at first, when the paths were of equal length. The difference between the lengths of the two tubes when interference occurs is equal to half the wave-length of the note used.

The interference of sound-waves is also very clearly shown in the case of the waves produced in air by a tuning-fork. As we shall see later, the extremities of the two prongs shown in section at A and B (Fig. 237) vibrate in such a way that they move alternately away from and towards each other. As a result, while they produce a condensation in the air between, they produce a rarefaction in the air towards E and F, and *vice versa*. Hence each prong starts two sets of waves, and

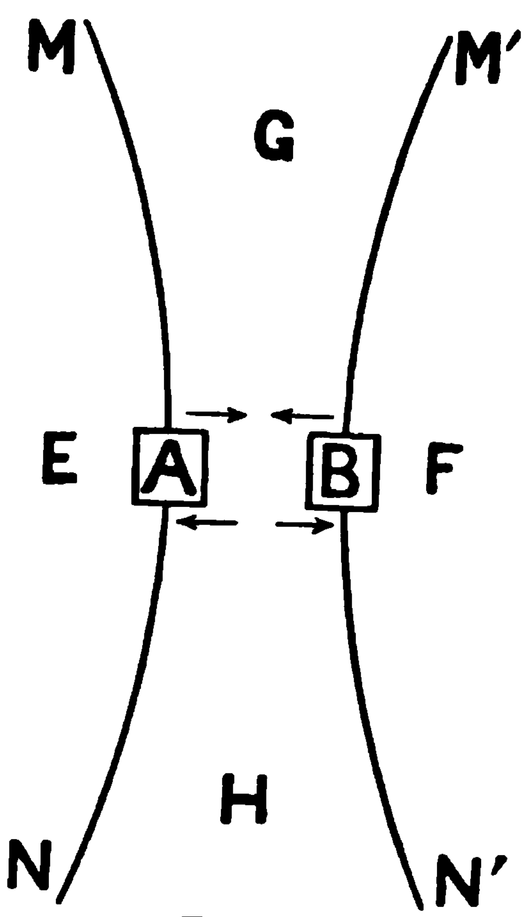


FIG. 237.

these waves are in opposite phase on the two sides, so that they interfere along the surfaces MAN and M′BN′. This interference can be very clearly distinguished by sounding a fork and then turning it slowly when held near the ear. When the ear is in either of the positions E, F, G, or H, the fork will be heard distinctly, while when the ear is on either of the curves MAN or M′BN′ no sound will be heard.

295. Stationary Waves formed by Reflection in Free Air.—When a shrill whistle or a bird-call P (Fig. 238) is sounded in front of a reflecting surface AB, the sound-waves which fall on the surface are reflected, and the reflected and incident waves interfere. Then, as in the case of the water-waves dealt with in § 275, there will be interference at the points N_1, N_2, N_3, &c., when the distance MN_1, MN_2, MN_3, &c., are equal to $\lambda/4$, $3\lambda/4$, $5\lambda/4$, &c., respectively. If then a sensitive flame is moved along the normal PM, it will not flare when it is at either of the nodes N_1, N_2, N_3, &c. Hence, since the distance between each of these points is equal to half the wave-length (λ), by noting the positions of the flame when it does not flare we may determine the wave-length of the note given by the whistle, and from the wave-length, knowing the velocity of sound in air, calculate the frequency of the note. As a sensitive flame will detect the presence of sound-waves of frequency too great to be heard by the ear, this method allows the frequency of such inaudible sounds being determined.

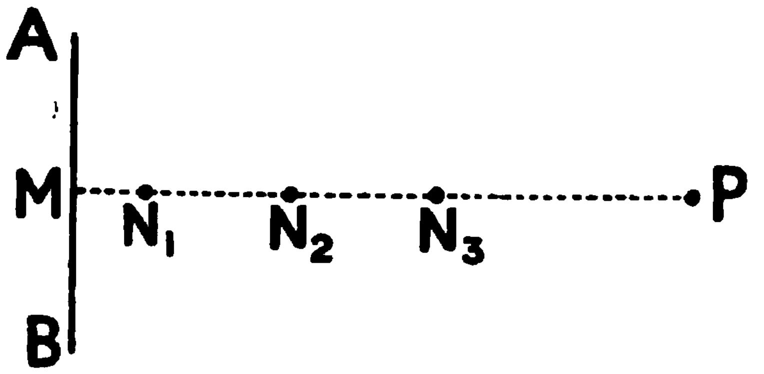


FIG. 238.

296. Dopler's Principle.—Suppose that at a point A there is a body which is emitting a note, of which the frequency is n. Owing to the action of the sounding body there will be a succession of waves produced in the surrounding air, and the frequency of these waves will also be n. Hence, if an observer is at a point B, n waves will reach his ear in each second, and he will hear a note of pitch n. Now suppose the observer approaches the sounding body, then in each second he will now receive more than n waves, for in addition to the n waves which would reach his ear, suppose he had been stationary, he will have met a certain number of waves in each second, for at the end of the second he is nearer the sounding body than he was at the commencement, and his ear will have received the waves which, at the commencement of the second, occupied this space. The result is that as he now receives more than n waves per second, the pitch of the note he hears will be higher than n. If, instead of approaching the sounding body, he moves away from it, then in the same way the pitch of the note heard will appear lower than n. If the observer remains at rest, but the sounding body approaches, then the effect will be the same as if the observer approached a stationary

sounding body, and the note heard will be of a higher pitch than that produced by the body; while if the sounding body is receding, the note heard will be lower.

This apparent change of pitch, owing to the relative motion of the sounding body and the observer, is called the Dopler effect, and the explanation which we have given is called the Dopler principle.

The Dopler effect can often be observed in the case of the note given by a steam-whistle sounding on an engine which is passing through a station at a rapid rate. The note heard when the engine is approaching the observer is very markedly of a higher pitch than that heard when the engine has passed and is travelling away from the observer.

The change in pitch produced by the relative motion of the observer and sounding body can readily be calculated. Suppose that the sounding body has a frequency of n vibrations per second, and that the observer is approaching it with a velocity v, the velocity of the sound being V. If the observer were at rest, he would receive n waves per second. If λ is the wave-length of the sound, then in a space v there will be v/λ waves. Hence, since the observer traverses a space v in one second, his ear will pick up, owing to his motion alone, v/λ waves, so that the total number of waves received in a second will be $n+v/\lambda$. But $V=n\lambda$ (§ 267), so that the pitch of the note heard is $n+nv/V$, or $n(1+v/V)$. If the observer had been travelling away from the sounding body, then in each second his ear would have failed to pick up v/λ waves, and the pitch of the note heard would be $n(1-v/V)$.

We may arrive at the same result by a rather different method of argument, which is instructive from its bearing on some other problems.

Suppose O_1, O_2, O_3, &c. (Fig. 239), to be the positions of the sounding body at the times o, t, $2t$, $3t$, &c. Then if, as before, V is the velocity of sound, and we describe a circle with O_1 as centre, and radius $4tV$, this circle will represent the wave-surface at a time $4t$ for a disturbance which was caused at the time o, when the sounding body was at O_1. In the same way if, with O_2, O_3, O_4 as centres, circles are described of radii $3tV$, $2tV$ and tV respectively, these will each represent the positions of the wave-surfaces at the time $4t$, that is, when the sounding body is at O_5, due to the disturbances produced when the sounding body was at the points O_2, O_3, and O_4 respectively. The figure very clearly shows the crowding together of the waves on one side, corresponding to a rise in pitch, and the spreading out in another direction, corresponding to a fall in pitch. The figure is drawn for the case where the velocity of the sounding body is less than that of the sound, the corresponding figure for the case where v is greater than V is shown in Fig. 240. The fundamental difference between the two cases is that when the velocity of the sounding body is less than that of sound, the wave-surface corresponding to the disturbance that starts at any given time lies entirely outside all the wave-surfaces corresponding to disturbances that start at all

subsequent times, while in the other case the whole or part of the wave-surface corresponding to subsequent disturbances lies outside. Thus

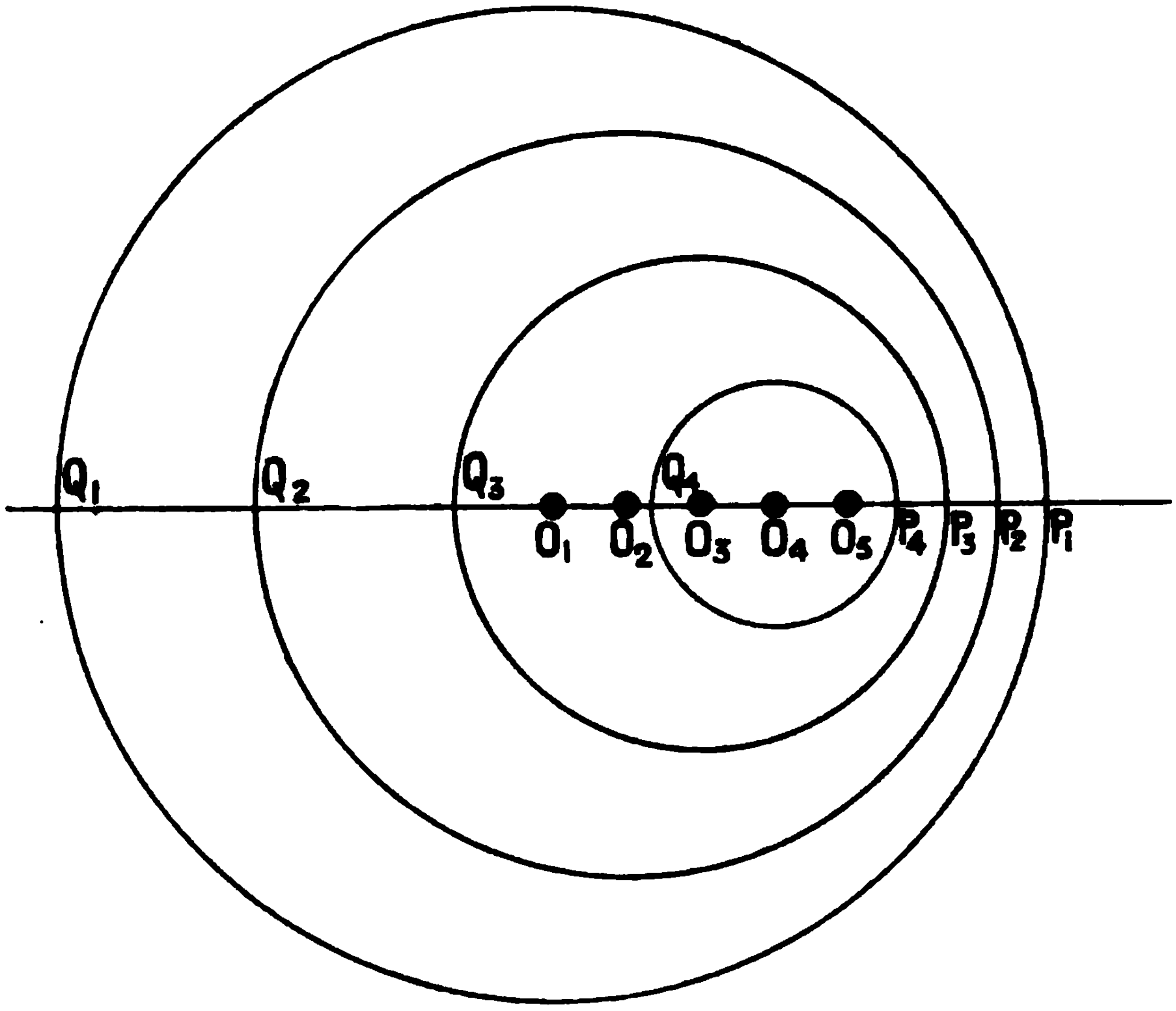


FIG. 239.

when the velocity of the sounding body is greater than the velocity of sound, a tangent can be drawn to the wave-surfaces at any given instant.

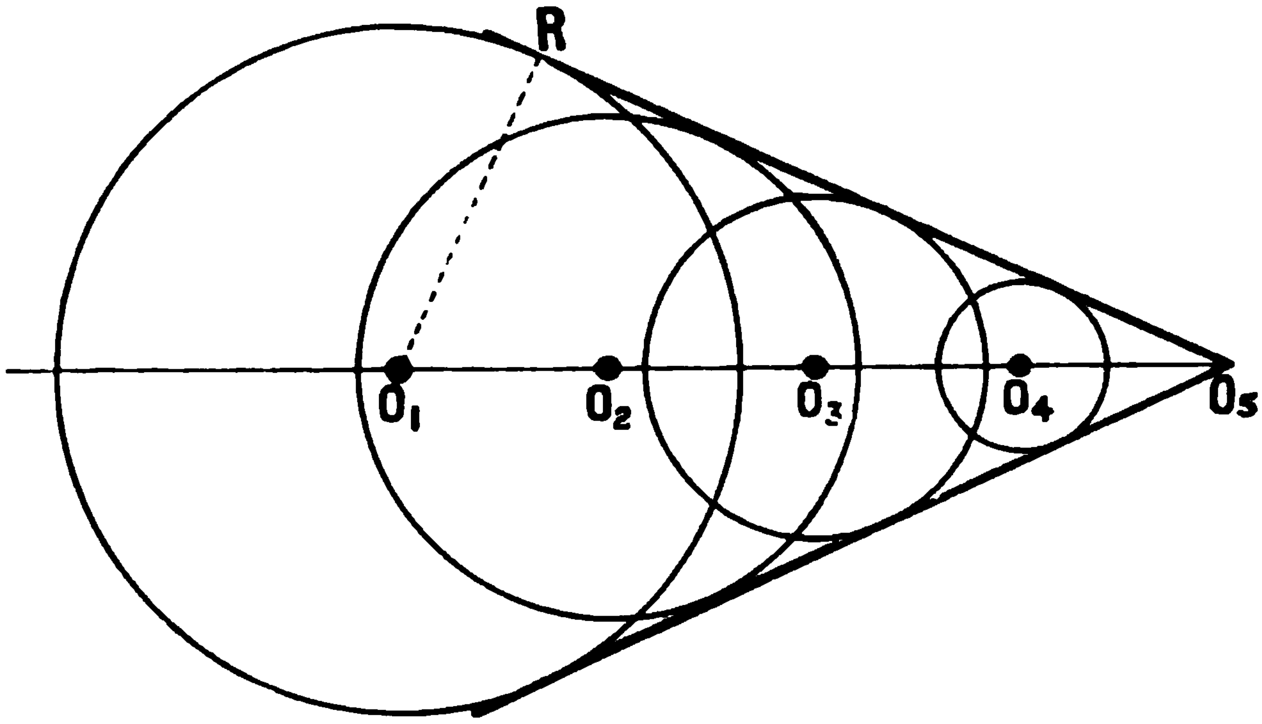


FIG. 240.

Since there are an infinite number of wave-surfaces intermediate to those shown in the figure, and all these wave-surfaces will have the same tan-

gents, and that, just as in the case considered in § 273, these waves will strengthen each other along these tangents but will interfere at all other points, two plane waves inclined at a constant angle will be produced. Since the moving body will reach the point O_6 at the same instant that the sound-wave reaches the point R, it follows that if θ is the angle made by the wave-front RO_6 with the direction of motion of the sounding body, then

$$\sin\theta = \frac{O_1R}{O_1O_6} = \frac{4tV}{4tv} = \frac{V}{v}.$$

Hence the greater the velocity v of the sounding body, the smaller the angle included between the wave-fronts. If we can measure this angle and know either the velocity of the moving body or that of sound, then the velocity of sound, or of the moving body, as the case may be, can be calculated.

A familiar example of this phenomenon, in the case of water-waves, is the bow-wave, produced when a boat moves through still water. Since the condition for the production of these waves is that the disturbing or sounding body must move faster than the waves themselves travel, it follows that in the case of sound-waves the velocity must exceed about 333 metres, or 1090 feet per second. In Fig. 241 the form of the wave produced in air by a rifle bullet is shown, as obtained in the photographs taken by Prof. C. V. Boys.

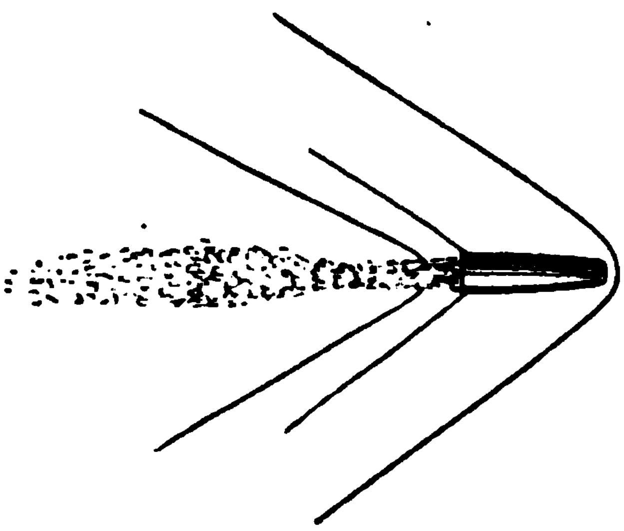


FIG. 241.

CHAPTER V

VIBRATIONS OF STRINGS, RODS, PLATES, AND COLUMNS OF GAS

297. Vibrations of Strings.—For our purpose we shall take a string to be a perfectly flexible, uniform filament of matter, stretching between two points. Any real string will, of course, possess rigidity; since, however, by taking a string of which the diameter is very small compared with the length, the effects of rigidity are very small, the errors thus introduced will not be great.

Strings can vibrate in two distinct manners, according as the vibrations of the particles which compose the string are longitudinal or transverse. We shall at present confine ourselves to transverse vibrations, postponing the consideration of longitudinal vibrations till we are dealing with the vibrations of rods.

Transverse vibrations may be produced in strings by plucking, as in the case of a guitar, bowing, as in the violin, or striking, as in the case of the piano.

The instrument usually employed for the study of the transverse vibration of strings is shown in Fig. 242, and is called a monochord or

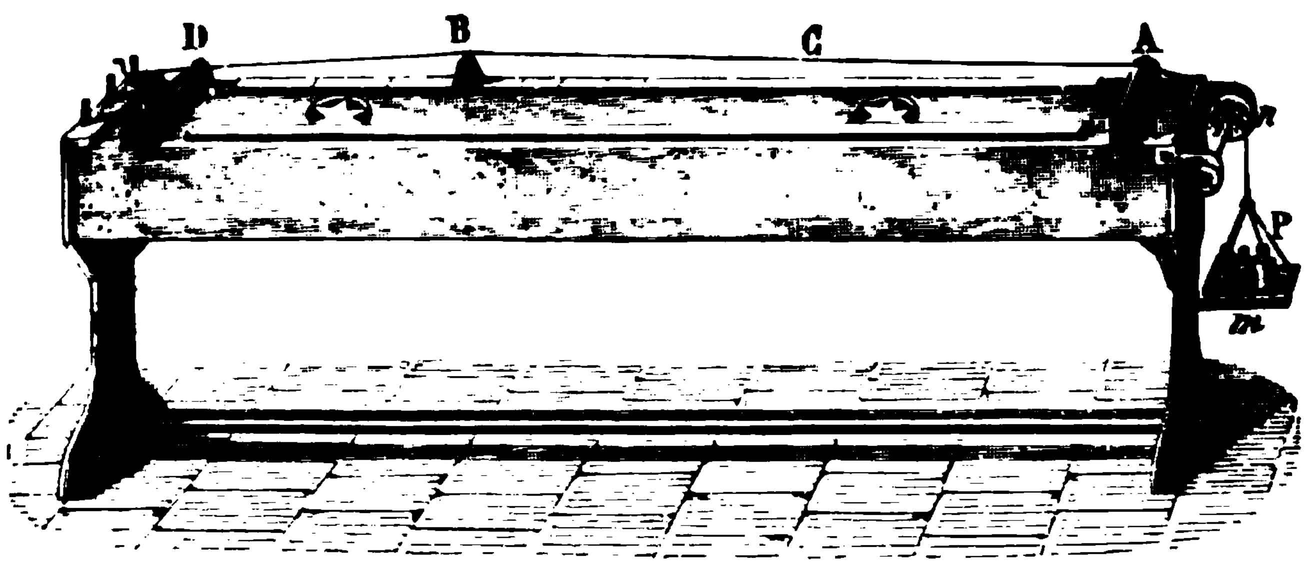


FIG. 242.
(From Ganot's "Physics.")

sonometer. It consists of a sounding-box, across the top of which a string is stretched by some weights P. In addition to the fixed bridges, A and D, there is a movable bridge, B, by means of which the length of the portion of the string which is put in vibration can be altered.

If a stretched string is struck or otherwise distorted and then let go, the velocity with which the wave set up travels along the string is given by the expression (§ 276)—

$$v = \sqrt{T/m} \quad . \quad . \quad . \quad (1),$$

in which T is the stretching force in dynes and m is the mass of one centimetre of the string in grams.

If a wave travelling along a string AB (Fig. 243) reach a point B, where the string is fixed, there will be reflection.

There is a difference between the reflection in this case and that which occurs when a water-wave or sound-wave in air is reflected from a solid obstacle. In the case of the water-wave, when a crest reaches the reflector it is reflected as a crest, while when a trough reaches the obstacle a trough is reflected. In the case of a crest travelling along a stretched string and reaching the *fixed* end of the string, it is not reflected as a crest but as a trough. Thus if A (Fig. 243) represents a transverse wave travelling up to the fixed end B of the string, then after reflection the form of the wave is that shown at C. The easiest way of seeing why this should be, is to suppose that the incident wave which moves up to the fixed end B of the cord (Fig. 244) is not reflected, but moves on past B along an imaginary continuation, BC, of the cord, and that the corresponding reflected wave is supposed to move on to the real part, BA, of the cord from the imaginary part CB. Now it is evident that the condition which the reflected wave has to fulfil is, since the point B on the string is held fixed, that the algebraical sum of the displacements at the point B on the string due to the direct and reflected waves must always be zero. At (*a*) (Fig. 244) the incident wave is shown at the instant when it has just reached the fixed point B. At (*b*) the incident wave, as shown by the light continuous line, has partly moved on to the imaginary portion of the cord, while the reflected wave, as shown by the dotted curve, has partly moved on to the real part of the cord. Under the influence of the direct and the reflected waves, the cord takes the position indicated by the thick line. The subsequent state of affairs is shown at (*c*), (*d*), (*e*), &c., and finally at (*i*) the incident wave has passed wholly on to the imaginary part of the cord, while the reflected wave consists of a wave moving to the right, but with the depression leading, although the incident wave had the elevation leading. This change in the condition of the wave due to reflection is equivalent to a loss or gain of phase equal to half a wavelength at the point of reflection, and so this type of reflection is often referred to as reflection *with* change of sign.

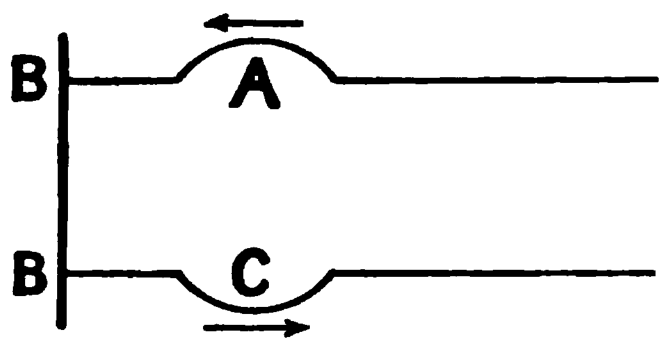


FIG. 243.

If in place of a single incident wave we have a train of waves, then, owing to interference between the incident and reflected waves, stationary waves will be formed. The point of reflection will in this case, however, be a node, and not a loop, as was the case with the water-waves considered in § 275. Fig. 228 will, however, apply if we suppose the point N_1 to be the fixed end of the string.

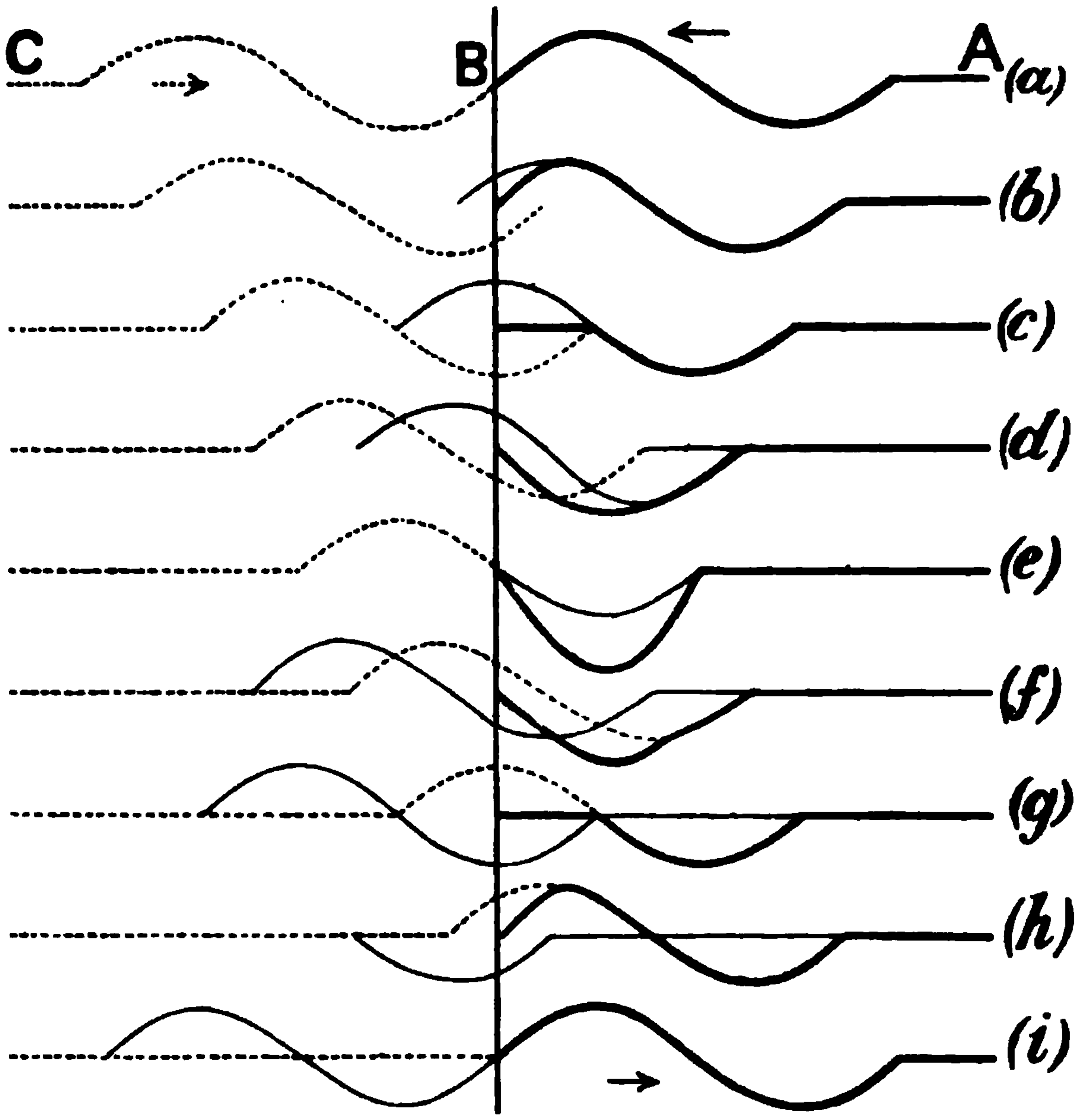


FIG. 244.

There will be a series of nodes at distances $\lambda/2$ from one another, and half-way between each node will be a loop.

Since the nodes are points where the string is permanently at rest, the vibrations will not be affected if we clamp the string at any of these points, in which case we should be dealing with the vibrations of a string held at each end.

If the second fixed point is placed at the first node from the fixed end, then the string will be vibrating so as to give the lowest tone of which it is capable, or, in other words, will be giving its fundamental. If l is the length of the string, we have $l=\overline{N_1N_2}=\frac{\lambda}{2}$. Now if n is the frequency of the string, we have, from the general equation connecting n, λ, and v (see § 267)—

$$v=n\lambda.$$

Hence from equation (1) above

$$n\lambda=\sqrt{\frac{T}{m}},$$

or, substituting for λ in terms of l,

$$n=\frac{1}{2l}\sqrt{\frac{T}{m}} \quad . \quad . \quad (2).$$

This expression may be written in a slightly different form, for if ρ is the density of the material of which the string is composed, and r is the radius of the string, $m=\pi r^2\rho$. Hence

$$n=\frac{1}{2l}\sqrt{\frac{T}{\pi r^2\rho}}$$

or

$$n=\frac{1}{2}\sqrt{\frac{T}{\pi r^2\rho l\,.\,l}}.$$

Now $\pi r^2\rho l$ is the total mass of the string, so that if we call this quantity M, the formula reduces to

$$n=\frac{1}{2}\sqrt{\frac{T}{lM}} \quad . \quad . \quad . \quad (3).$$

In addition to the fundamental mode of vibration considered above, the string can, as shown in Fig. 245, vibrate in such a way that there are 1, 2, 3, &c., points which are permanently at rest between the extreme points. The points N_1, N_2, &c., which are permanently at rest during the vibrations, are nodes, while the points marked L, at which the amplitude of the vibrations is a maximum, are loops.

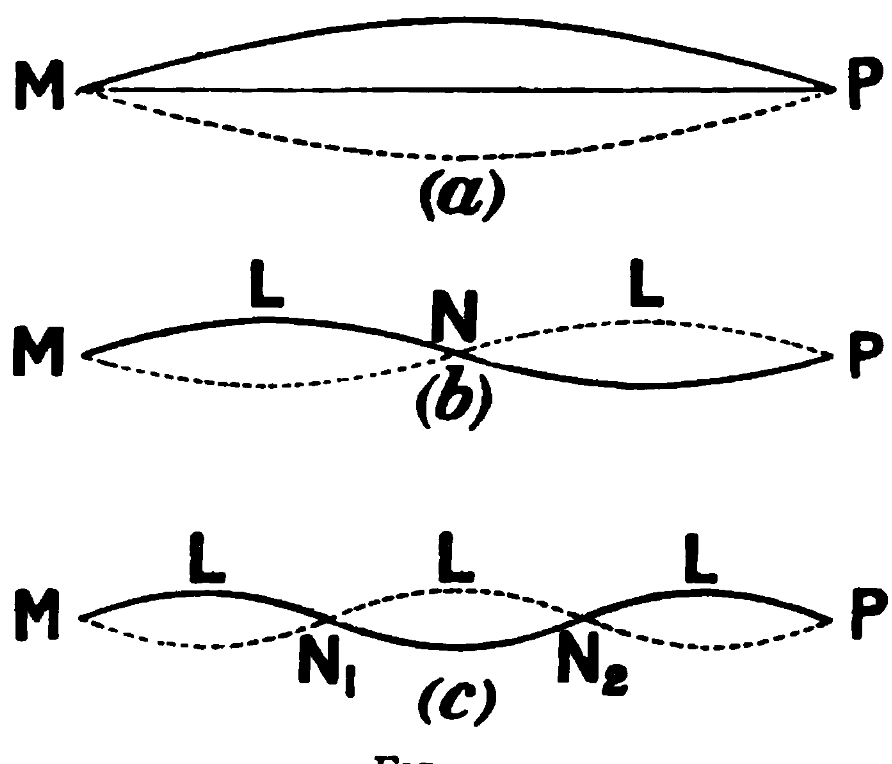


FIG. 245.

Since if the string in (*b*) (Fig. 245) were clamped at N the vibrations would be unaffected, and under these circumstances we should be dealing with two strings each of which had a length $l/2$, we see that the frequency of the tone produced under these conditions will be obtained by substituting $l/2$ for l in the equation (2), since the tension, T, and the mass per unit length, m, remain the same. Hence the frequency n' is given by

$$n'=\frac{1}{l}\sqrt{\frac{T}{m}},$$

that is, the frequency is double what it was in the case when the string

was vibrating as a whole. In this way it will be seen that, if n is the frequency of the fundamental tone of a string, then the overtones have the frequencies $2n$, $3n$, $4n$, &c. The string can give, therefore, the harmonics of the fundamental tone.

We have supposed above that the string vibrates in only one of the possible ways at a time; we may, however, have a number of them coexisting, so that a note will be produced, which is built up of the tone and certain of the overtones.

Thus, if the string is plucked at a third of its length from either end the first overtone will be particularly noticeable, and can be rendered even more evident if the fundamental note is damped out by lightly touching the string at its middle point with a feather or piece of paper, when the octave of the fundamental, which corresponds to the first overtone, will be heard very clearly.

298. Melde's Experiment.—A very convenient way of showing the laws which govern the vibrations of strings is that due to Melde. A string AB (Fig. 246) is fixed at one end to the prong of a large tuning-fork, while the other end passes over a pulley and has a small scale-pan P attached, by putting weights in which the tension on the string can be varied at will. If the fork is placed in the position shown in the figure, then when the prong B moves towards A the string will be allowed to sag. As the prong moves back the string is tightened, being horizontal when the prong is furthest from A. When the prong moves back towards A the string does not again sag, but, owing to its inertia and the upward motion it has acquired, it travels up above the horizontal position, so that it is in its extreme upward position when the prong has completed one whole vibration, from the time when the string was at its lowest. Thus the frequency of the string is half that of the fork.

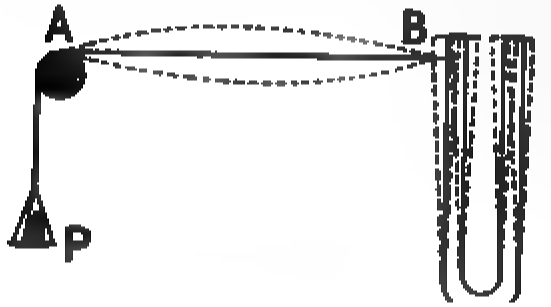


FIG. 246.

If the fork is turned through 90°, so that the movement of the prong is at right angles to the string, then when the prong moves to the right the cord will also move to the right, when the prong is passing through its position of rest the string will also be passing through its position of rest, and when the prong is at its extreme left elongation so also will be the string. The string therefore will in this case vibrate in unison with the fork. In order to obtain a considerable movement of the string, it is necessary that the tension should be adjusted so that the natural period of the string should agree with the period of the vibrations impressed on it by the fork.

If the fork is in the position shown in Fig. 246, and has a frequency N, then when the tension T_1 of the string is such that

$$\tfrac{1}{2}N = \frac{1}{2l}\sqrt{\frac{T_1}{m}},$$

the vibrations set up in the string will be most energetic, the string vibrating in its fundamental mode. As the tension is decreased the amplitude of the vibrations will decrease, but when the tension reaches the value T_2 given by the equation

$$\tfrac{1}{2}N = \frac{1}{2l/2}\sqrt{\frac{T_2}{m}}$$

the string will again vibrate strongly, but now with a node in the middle. By further decreasing the tension 3, 4, 5, &c., nodes can be produced, but for the vibrations to be energetic the tension, T, has in every case to be so adjusted that if l' is the distance between consecutive nodes,

$$\tfrac{1}{2}N = \frac{1}{2l'}\sqrt{\frac{T}{m}}.$$

299. Transverse Vibrations of Rods.—In considering the vibrations of strings, we have neglected the rigidity of the string, and supposed that the only force tending to bring the string back into its position of rest when it is displaced is that due to the tension. We now pass on to the opposite extreme, in which the restitution is due solely to the rigidity of the solid, namely, the case of a rod clamped at one or more points but not subjected to any tension.

If the rod is held in a clamp at one end, the fundamental form in which it can vibrate is shown at (*a*), Fig. 247, where there is a single node, and that at the fixed end.

The manner in which a rod clamped at one end vibrates, when sounding its first and second overtones, is shown at (*b*) and (*c*). If the frequency of the rod when sounding its fundamental is taken as unity, then the frequencies of the overtones are 5.29, 8.27, 10.21, 11.66, &c. In this case it will be observed that the overtones are not the harmonics of the fundamental tone.

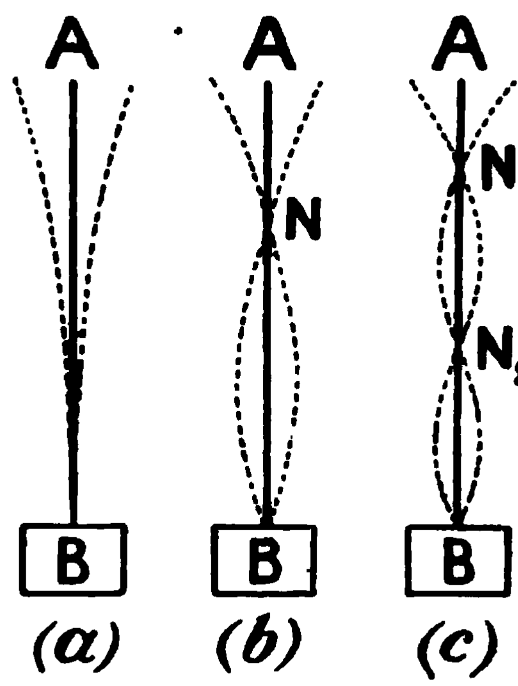


FIG. 247.

If the rod, when vibrating as in (*b*), Fig. 247, instead of being clamped at B, were prolonged, and were simply supported at N and B, we should have the case of a rod free at both ends, and vibrating in its fundamental form as shown at (*a*), Fig. 248. The mode

of vibration for the first overtone is shown at (*b*), and is such that there are three nodes. The relative frequencies of the fundamental and of the overtones for a rod free at both ends are 1, 2.92, 4.87, 6.32, 7.48, &c.

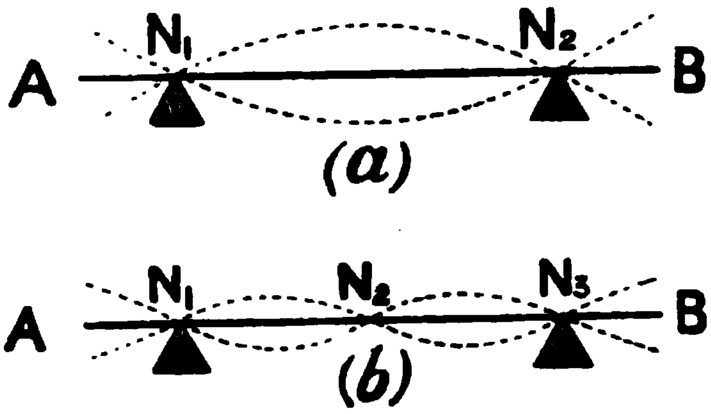


FIG. 248.

The consideration of the connection between the dimensions of a rod and the frequency of its fundamental tone is beyond the scope of this work. It is, however, interesting to note that, other things being the same, the frequency varies inversely as the square of the length, and in the case of rectangular rods, directly as the thickness, or as the radius in cylindrical rods.

300. Tuning-Forks.—If we consider the vibrations of a rod, free at both ends, when sounding its fundamental, so that there are two nodes in the positions shown in (*a*), Fig. 249, and suppose that the rod is gradually bent at the middle, it is then found that the nodes come nearer together as the rod is bent, as shown at (*b*), (*c*), and (*a*) in the figure. When the two limbs are parallel the arrangement forms a tuning-fork, and the nodes practically coincide. For convenience in holding the fork a stem is attached at the middle point, and since this point is a node, the vibrations are not interfered with on this account.

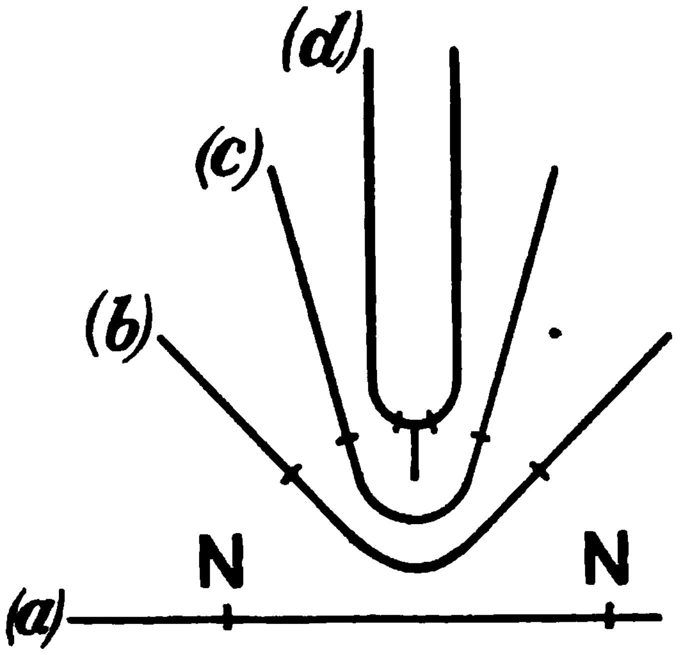


FIG. 249.

The above method of considering the tuning-fork, as derived from a straight rod, shows how it is that the prongs vibrate in such a way that the ends alternately approach and recede from one another, for in the straight rod, where there is a loop in the middle, the ends move up and down together, and this continues in the bent rod, although the central loop has now vanished.

The reason that tuning-forks are of such importance in sound is that the note given by a properly formed fork, when set in vibration by gentle bowing, is practically a simple tone, for the overtones are comparatively difficult to obtain, and, when present, die out very much more quickly than the fundamental tone. The absolute frequency of the tone given by a fork depends on the temperature of the fork, owing to the change of the elasticity of the metal of which the fork is com-

posed. In the case of a steel fork, the frequency decreases by about 1 in 8943 for each degree Centigrade rise of temperature.

Since tuning-forks are almost exclusively used for standards of pitch, it is of importance to be able to determine the pitch of a fork with accuracy.

One method by which the frequency of a fork can be measured is to tune a string, which is stretched by a known force, to unison with the fork, and to calculate the pitch of the note given by the string by the formula given in § 297. This method does not, however, admit of any great accuracy. A better method consists in attaching a fine bristle to one prong of the fork, and causing this bristle to trace a wavy line on a smoked drum, which is rotated at a fairly rapid rate. A small electro-magnet works another style, which traces a second line on the drum alongside that due to the fork. The current of the electro-magnet is made at equal intervals by the pendulum of a standard clock, so that the line traced on the drum is broken, and the time interval between these breaks is known. Hence, by counting the number of vibrations of the fork which occur between two given

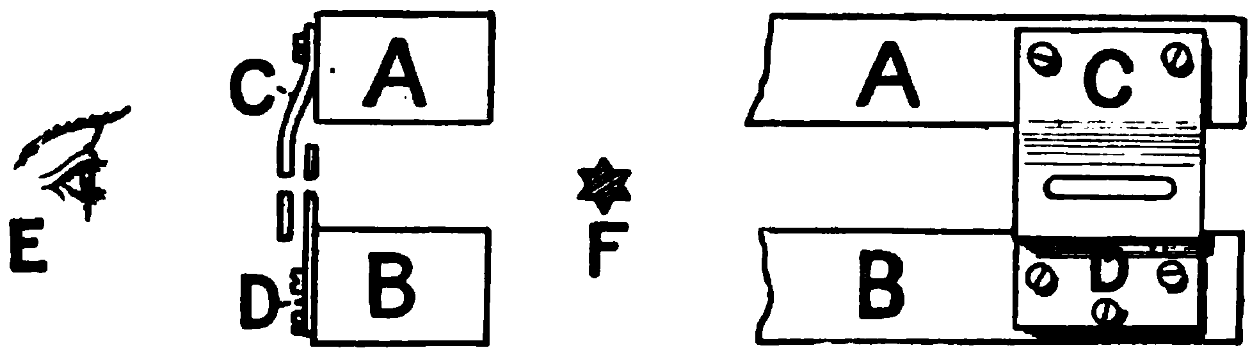


FIG. 250.

time-marks, the frequency of the fork can be obtained. Another method of considerable importance is that known as the stroboscopic disc method. Two thin and light pieces of card or metal foil C, D (Fig. 250), are attached to the prongs A, B of the fork. Each of these cards is perforated by a slit, and these slits are so placed that, when the prongs of the fork are at rest, the two slits are opposite each other, so that an eye placed at E can see through. When the fork is sounding, it will only be possible to see through when the prongs are passing through their position of rest, and hence an object placed at F will be seen intermittently, the interval between two views being equal to the interval between two consecutive passages of the prongs through their positions of rest, that is, at intervals equal to half the period of the fork. If the object at F is a disc on the face of which are painted a number of rings of equidistant dots, as shown in Fig. 251, then, if this disc is in rotation, and, during the time which elapses between two views through the slits, a dot in any one of the rings has just had time to take the position occupied by the preceding dot,

when the disc was seen before, this ring of dots will appear at rest, for the eye cannot distinguish between the different dots, and whenever it sees the disc, the dots on the ring considered are as a whole in the same position. Let the velocity of rotation of the disc be gradually increased till the dots in one of the rings, when viewed through the slits, appear at rest, and suppose that the angle subtended by two adjacent dots of this ring at the centre of the disc is θ, so that if there are m dots in the ring, $\theta = 360/m$, and that the disc makes n turns in a second, so that the time taken to make one turn is $1/n$. Then the time taken to turn through the angle θ is $\theta/360n$ or $1/mn$. Hence the time taken by the fork to make half a vibration is $1/mn$, or the frequency is $2mn$.[1] The quantity n is obtained by counting the number of turns made in a given time by means of an arrangement similar to that shown attached to the syren in Fig. 233, the speed of rotation being kept constant during this time by observing that the ring of dots appears at rest throughout.

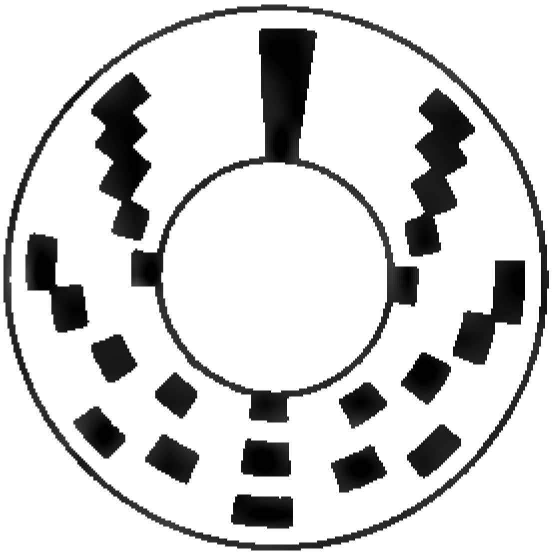


FIG. 251.

301. Lissajous' Figures.—Since the movement of the prong of a tuning-fork is a simple harmonic motion, it is possible by means of two forks to illustrate the combination of two simple harmonic motions treated of in sections 53, 54. If a mirror M (Fig. 252) is attached to the prong of a fork A, and a ray of light is reflected from the mirror and received on a screen S, then, when the fork is set in vibration, instead of getting a spot of light on the screen, we shall get a line, owing to the motion of the mirror and to the persistence of vision. The line will be parallel to the limb of the fork, that is, in the case shown in the figure it will be vertical. If now the fork is rotated round a

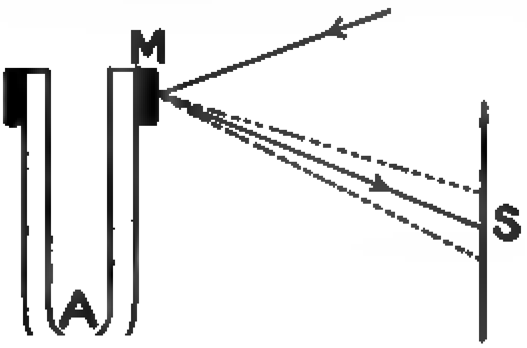


FIG. 252.

[1] This result may be obtained otherwise, thus. Each time the disc is seen, each dot has moved on one. Now the number of dots which pass any given point in a second is nm. Hence the disc must be seen nm times per second. But it is seen $N/2$ times per second, where N is the frequency of the fork. Therefore $N/2 = nm$.

vertical axis, a wavy curve will be produced on the screen, resembling a sine curve, for, owing to the rotation alone, we should obtain a horizontal straight line, so that when the fork is moving we have to compound a steady horizontal motion with a S.H.M. in a vertical direction.

Next let the ray of light, after being reflected by the mirror attached to the fork A, be again reflected at a mirror attached to a second fork B (Fig. 253), this fork being placed so that the motion of its prongs takes place in a direction at right angles to that of the prongs of A. If A alone vibrated, the line aa' would be produced on a screen S; while if B alone vibrated, the line bb' would be produced. When they both vibrate together, the spot of light will trace out a curve due to the combination of the two motions. If the frequencies of the forks are in the ratio of 2 to 3, the curve will be one of those shown in Fig. 40. If one fork is the octave of the other, the curve obtained will be one of those shown in Fig. 41. Which of the forms is obtained with any given ratio depends on the relation between the phases of the two forks. If the frequencies are exactly commensurable, the form of the curve obtained will remain constant. If, however, the ratio of the frequencies is not quite in the ratio of two simple numbers, say they are very nearly the octave, then, as shown in § 53, the curve will gradually change, taking in turn all the forms shown in Fig. 41.

FIG. 253.

The passage of the curve through the different forms belonging to any given ratio of the frequencies, when the forks are not quite adjusted to this ratio, is a very accurate method of telling by how much the ratio of the frequencies differs from the correct ratio, or of adjusting them so as to give the correct ratio, for, as was shown in § 53, all the forms will be gone through once while one fork gains a vibration on the other. Thus if m and $n+\delta$ are the frequencies of the two forks, where δ is a small quantity, so that the figures obtained correspond to the ratio $m : n$, the whole series of curves corresponding to this ratio will be gone through in a time t, such that $t\delta$ is equal to unity. Hence if we note t we can calculate δ, and therefore deduce the true ratio, $m : n+\delta$, of the frequencies of the two forks.

The changes which take place in the form of Lissajous' figures being such an accurate method of adjusting the frequencies of two forks to certain fixed ratios, it is of much use in adjusting the pitch of forks. Since, however, ordinary forks are not fitted with a mirror, and the addition of a mirror would alter the pitch, the arrangement described above is not applicable. Hence the arrangement shown in Fig. 254, and called a vibration microscope, is employed. A large fork A, which is the standard with which the others are to be compared, carries attached to one of its prongs a small lens B. This lens forms the objective of a small microscope, C, the tube and eye-piece of which are supported on a separate stand. The fork D, which is being adjusted, is placed so that its prongs vibrate in a direction at right angles to those of the standard A. If now the microscope is focussed on a small dot on the top of one of the prongs of the fork D, and the standard fork is alone sounding, the dot will appear drawn out into a line parallel to the line F, owing to the to-and-fro motion of the lens B. If, on the other hand, the fork A is at rest, and the fork D is sounding, the dot will appear as a line parallel to the line E. When both forks are sounding, the pattern traced out by the dot will be the Lissajous' figure appropriate to the relative frequencies of the forks. Since in this method of obtaining Lissajous' figures no addition has to be made to the fork which is being tested, it can be used for adjusting any form of fork.

FIG. 254

The pitch of a tuning-fork is adjusted by filing the prongs. If some of the metal is filed away near the extremity of the prongs, the pitch will be raised, for the mass of metal which swings backwards and forwards is thus decreased. If the filing is performed near the stem the pitch is lowered, for this decreases the stiffness of the prongs, and so reduces the restoring force which acts when the prongs are deflected from their position of rest.

302. Transverse Vibrations of Plates.—If a square plate of brass or glass is supported at its centre in a horizontal plane by being screwed to a vertical pillar, and the edge is bowed with a rosined fiddle-bow, it will give out a note. The number of notes which can be obtained from any one plate is, however, practically infinite. By strewing sand on the upper surface the character of the vibration can be studied, since the sand gathers along the nodal lines,

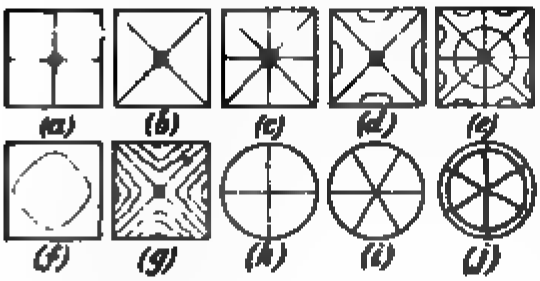


FIG. 255.

i.e. the lines where the plate remains permanently at rest. Some of the forms which can thus be obtained are shown in Fig. 255, which are selected from some of those published by Chladni, and the figures obtained in this way are called Chladni's figures.

The form shown at (*a*), Fig. 255, is obtained by damping the plate, by touching it with the finger, at the middle of one of the edges, and bowing at one of the corners. To obtain (*b*) the corner is damped and the plate bowed in the middle of an edge, and so on, the bowing being always performed at the point half-way between two nodal lines, and the damping taking place at the nodal lines.

Wheatstone has given an elementary explanation of the forms in which a square plate ABCD (Fig. 256) can vibrate. He first considers the plate as made up of a number of rods parallel to one side, AB. These rods could vibrate so as to all have their nodes along the lines $N_1N'_1$, $N_2N'_2$. In the same way the plate may be considered as made up of rods parallel to AD, and these rods would all vibrate with their nodes along the lines $M_1M'_1$ and $M_2M'_2$. Now suppose the two movements to go on simultaneously, so that the actual motion of any portion of the plate is the algebraical sum of the movements due to the two sets of rods. Let us first take the case when the central segments of the two sets of rods are in opposite phase. Then, if we indicate that any given part of the plate is above the plane of the plate by the symbol +, and that it is below by −, we have the two separate motions indicated at (*a*) and (*b*), Fig. 257. When these two are combined we get the state shown in (*c*). In the rectangles which are shaded the two displacements assist one another, while in the unshaded portions the upward displacement due to one set of rods is neutralised by the downward displacement due to the other set. Hence the minimum displacement will take place along the two diagonals AC, BD, a result which agrees with the form shown at (*b*) in Fig. 255. If the central portions of the two sets of rods are in the same phase, then when they exist simultaneously the figure shown at (*f*), Fig. 257, will be produced. The nodal lines will therefore be the square EFGH. This case corresponds to (*f*) in Fig. 255; and since the centre of the plate is in vibration, the plate is not clamped at its centre, but at a point on the nodal line. It will be seen that the general form of the nodal lines obtained by experiment agrees fairly well with the form obtained by Wheatstone's method. The slight differences between the calculated and observed forms is to be expected, since we have assumed that the amplitude of the motion of the central portion of the

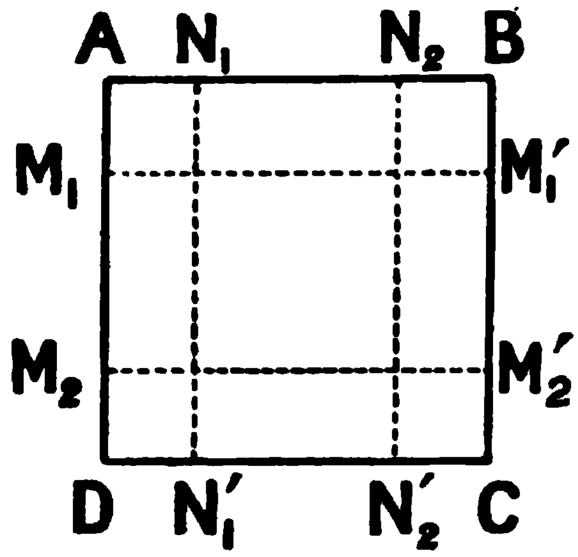


FIG. 256.

rods and of the extremities are equal. This, however, is not true, the amplitude of the ends being really considerably greater than that of the central segment. Hence some of the parts which we have taken as being at rest, owing to the displacement of the end segment of one set of rods neutralising the displacement of the central segment of the other set, will not really be completely at rest.

In the case of a circular plate, the nodal lines are either radial lines, or circles, or combinations of the two. The radial nodal lines are obtained by fixing the centre of the circular plate and bowing the edge, while two points on the edge are damped. Since a nodal line always represents the line of separation between two parts of the plate which are vibrating in opposite phases, there must always be an even number

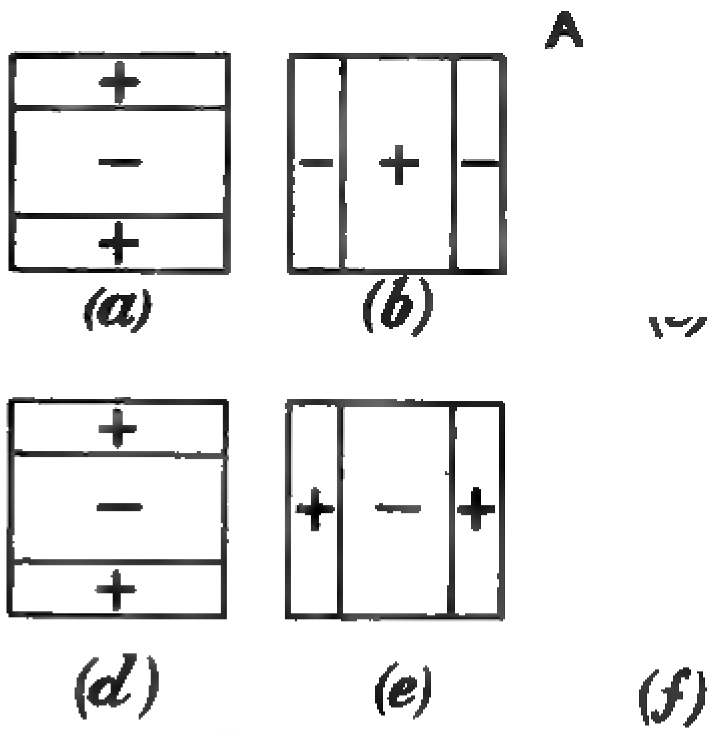


Fig. 257.

of radial nodal lines. If there were an odd number, then at one line at least the plate on both sides of the line would be vibrating in the same phase.

The circular nodal lines can be obtained by resting the plate on three points on one of the circles, and setting it into vibration by drawing a rosined string through a hole at the centre. They may also be obtained by fixing the plate by its centre to the end of a rod, and making this rod vibrate longitudinally (see § 304).

If, instead of using sand to show the nodal lines, a light powder, such as licopodium, is employed, it will collect, not along the nodal lines, but at the parts of maximum motion. This is due, as was shown by Faraday, to the formation of small vortices in the air near the plate, just above the

vertical segments or loops, these vortices sweeping the powder on to the loops.

303. Bells.—In the case of bells, as in the case of circular discs, there must always be an even number of nodal meridians, the portion of the bell on opposite sides of each meridian vibrating in opposite phase. The simplest form of vibration is that in which there are four nodal meridians, N_1, N_2, N_3, and N_4 (Fig. 258). Although the nodal meridians are points of no radial motion, they are points of maximum tangential motion. The reason is that when the rim on one side of a node is outside the mean position, the rim at the other side is within the mean position, and, as is obvious from the figure, the length of rim intercepted between adjacent nodes is greater when this portion of the rim is outside the mean position than when inside the mean position. The result is that, to allow for this change in the length of rim intercepted between adjacent nodes, a motion of the rim in its own plane takes place at the nodes.

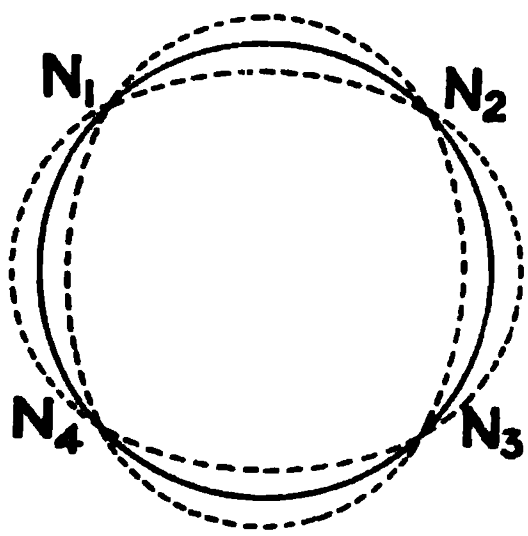


FIG. 258.

The overtones of a bell are not the harmonics of the fundamental, as is well shown by the following table, which gives the various tones of a peal of bells which were examined by Lord Rayleigh :—

Bell.	Nominal Pitch.[1]	Actual Pitch of Tones given by Bell.
5	*f*♯	*g* − 3,[2] *g*′ − 4, *a*′ + 6, *d*″ − 3, *f*″ ♯ − 2.
4	*g*♯	*a* + 3, *g*′ ♯ − 4, *b*′ + 6, *d*″ ♯ + 20, *g*″ ♯ − 6.
3	*a*♯	*a* ♯ + 3, *d*′ + 6, *c*″ ♯ + 4, *e*″ + 6, *a*″ ♯.
2	*b*	*d*′ − 6, *d*′ ♯ − 5, *d*″ + 8, *g*″ ♯ + 10, *b*″ + 2.
1	*c*♯	*d*′ + 2, *b*′ + 2, *c*″, *g*″ ♯ + 4, *c*‴ ♯ + 3.

An examination of this table will show how markedly inharmonic are the various overtones. The very curious fact also appears that in the case of all these bells, which are by English founders, it is the fifth of the above tones that fixes the nominal pitch of the bell.

304. Longitudinal Vibration of Rods and Strings.—In the case of a string or rod, in addition to the transverse vibrations already considered, we may have longitudinal vibrations, in which the particles of the string move backwards and forwards parallel to the length of the string.

[1] No attention is here paid to the question as to the *octave* of the nominal pitch of the bells.

[2] *g* − 3 means the frequency was three vibrations per second less than *g*.

The frequency of the longitudinal vibrations of a string is independent of the tension with which the string is stretched. For when a particle of the string is displaced from its mean position, the force with which it tends to return to its undisplaced position depends on the stress caused by the *displacement* from this position, and this stress is by Hook's law (§ 172) independent of any previously existent stress which affects the particle under consideration and the other particles equally. Thus the velocity with which a longitudinal disturbance travels in a string is independent of the tension, and depends only on the elasticity and density of the material of the string. When the string is giving its fundamental, there will be a node at each end and a single loop in between, so that the wave-length will be equal to twice the length of the string. Hence, since $v=n\lambda$, where v is the velocity of sound in the material of the string, n is the frequency of the note produced, and λ is the wave-length, we get $v=2nl$, where l is the length of the string.

In the case of a rod clamped at the middle and vibrating longitudinally, there must be a node at the centre where the rod is held, and the ends of the rod must always be loops. Hence when the rod is sounding its fundamental there will be a loop at either end, and a single node, namely that at the centre. The wave-length of the sound in the rod will therefore be equal to twice the length of the rod.

Since, as in the case of the longitudinal vibrations of a string, $v=2nl$, if we measure the frequency n of the note given by a rod or string of length l when vibrating longitudinally, we can immediately calculate the velocity of sound in the material of which the rod or string is composed, and it is in this way that the values for the velocity of sound in the solids given in the following table have been obtained.

VELOCITY OF SOUND IN SOLIDS.

	Cm./Sec.	Feet/Sec.
Aluminium	5104	16740
Brass	3500	11480
Steel	4990	16360
Glass	5000 to 6000	16410 to 19690
Pine (along the fibre)	3320	10900

We may also calculate the velocity of sound in a string or rod, and thus also calculate the pitch of the note given by such rod or string when vibrating longitudinally, for the waves concerned are of the type considered in § 279, and it was there shown that the velocity of such a wave is given by

$$v=\sqrt{E/\rho},$$

where E is the elasticity (Young's modulus (§ 172) in this case) and ρ is the density. Hence, since

$$n = v/2l,$$

we have

$$n = \frac{1}{2l}\sqrt{E/\rho}.$$

In the case of brass, Young's modulus has the value 1.1×10^{12}, and the density is 8.7. Hence the velocity of sound in brass is

$$v = \sqrt{\frac{1.1 \times 10^{12}}{8.7}} = 3550 \text{ cm./sec.},$$

a number which agrees with that obtained by experiment.

In the case of a rod clamped in the middle, the first overtone is produced by a mode of vibration in which there are three nodes, one of them being, of course, at the middle. The pitch of the note given is nearly three times that of the fundamental note. The next overtone contains five nodes, and the pitch corresponds to nearly five times that of the fundamental, and so on. If a rod is held with one end fixed and the other end free, there must be a node at the fixed end and a loop at the free end. Hence the wave-length of the fundamental, when such a rod is vibrating longitudinally, will be equal to four times the length of the rod, a result which follows immediately from the case of a rod clamped at the middle, for this latter may be regarded as made up of two rods clamped at one end. The positions of the nodes for the fundamental and the first two overtones are shown in Fig. 259, from which it will be immediately seen that the frequencies are as 1 : 3 : 5.

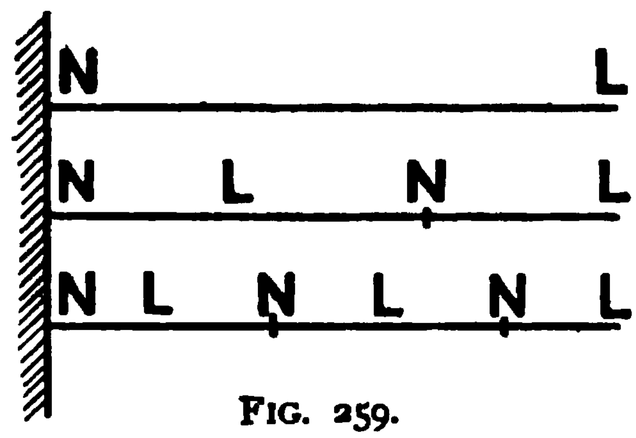


FIG. 259.

305. Torsional Vibrations.—When a rod is clamped at one end, and the side is bowed transversely with a rosined bow, a very high note can be obtained. The vibrations in this case consist of an alternate twisting and untwisting of the rod, and are called torsional vibrations. If the solid is in the form of a rectangular bar, and one face is held horizontal, by strewing sand on this face and bowing the edge of the rod it can be set in torsional as well as transverse vibrations, and the positions of the nodal lines will be shown by the sand.

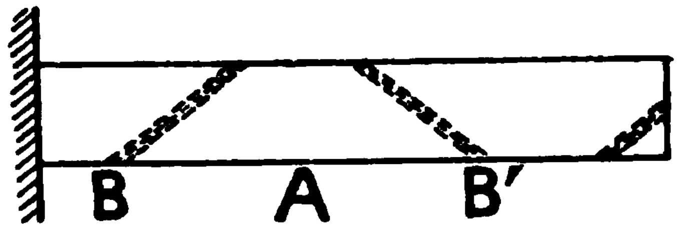


FIG. 260.

The character of the nodal lines thus obtained is shown in Fig. 260. If the vibrations were simply transverse, the nodal lines would be at right angles to the edge. Owing,

however, to the production of torsional vibrations, in addition to the transverse vibrations, when the bar is bowed at A and damped at B and B′, the nodal lines are inclined, as shown in the figure.

306. Vibrating Columns of Gas.—The column of gas, say air, enclosed in a tube can be caused to vibrate longitudinally in a manner strictly analogous to that of the longitudinal vibrations of rods. Two cases have to be considered, namely, that in which the tube is open at both ends, and that in which the tube is closed at one end.

In the case of vibrating columns of air, at the nodes, which are points where the air particles are at rest, there will be maximum change of pressure, for the particles will alternately be crowded together and separated at these points. The loops, on the other hand, will be places of maximum motion, but of minimum change of density and pressure.

In the case of a closed pipe, there can be no motion of the air particles which are in immediate contact with the closed end, so that the closed end must always be a node. At the open end, where the air column communicates with the external air, the changes of density can only be very small, so that the open end may for the present, at any rate, be regarded as a loop. Hence the fundamental is produced when the air column vibrates, as at (*a*), Fig. 261. The wave-length will be equal to four times the length of the pipe, for it is always equal to four times the distance between a node and the adjacent loop. The first overtone is produced when there is one node besides that at the closed end, as shown at (*b*), while the second overtone is produced when there are two additional nodes, as at (*c*). The air particles on the two sides of a node are always moving in opposite directions, and when a condensation is taking place at one node, a rarefaction is taking place at the adjacent nodes. The wave-length at (*b*) is equal to twice the distance between consecutive nodes, that is, is equal to $\frac{4}{3}l$, where l is the length of the pipe. Thus the wave-lengths of the fundamental and of the overtones of a closed pipe are—

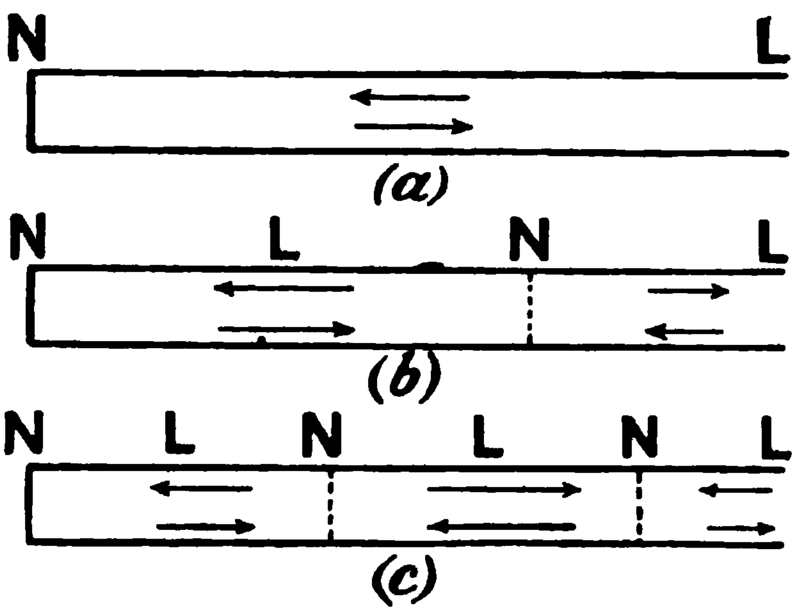


FIG. 261.

$$4l,\ \frac{4l}{3},\ \frac{4l}{5},\ \frac{4l}{7},\ \text{\&c.}$$

Since the velocity of sound in the air is the same in all cases, and $v = n\lambda$, the frequencies of the fundamental and of the overtones are inversely proportional to the wave-lengths, so that, if the frequency of the funda-

mental is taken as unity, the frequencies of the fundamental and of the overtones are 1, 3, 5, 7, &c. In this case, therefore, only the odd harmonics of the fundamental are present in the overtones.

In the case of a pipe open at both ends, there must be a loop at each end, and the fundamental is given when there is a single node produced at the middle, as shown at (*a*), Fig. 262. In this case the wave-length is equal to twice the length of the pipe. The modes of vibration corresponding to the first two overtones are shown at (*b*) and (*c*). It will be seen that the wave-lengths of the fundamental and overtones are equal to $2l$, $\frac{2l}{2}$, $\frac{2l}{3}$, $\frac{2l}{4}$, &c., or in the ratio of $1 : \frac{1}{2} : \frac{1}{3} : \frac{1}{4} : \ldots$ Hence the frequencies are in the ratio of

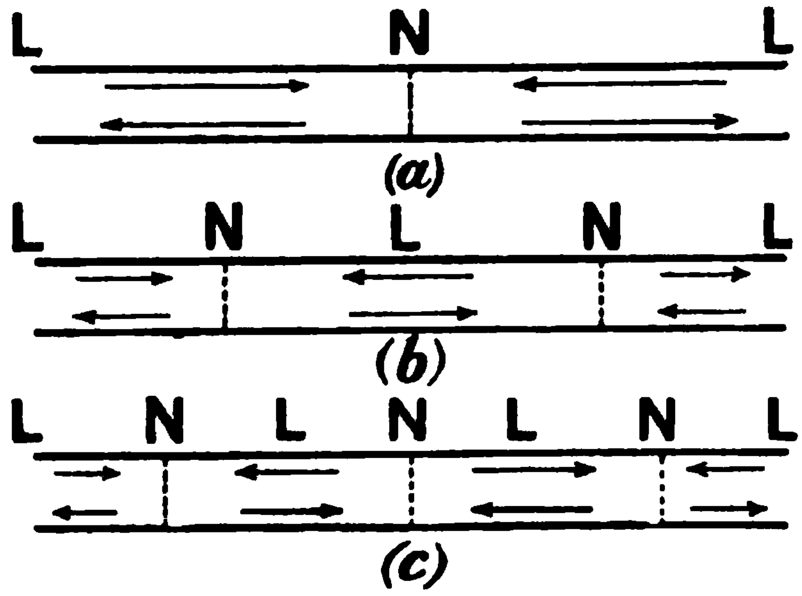


FIG. 262.

$$1 : 2 : 3 : 4 : \ldots$$

So that in the case of an open pipe all the harmonics of the fundamental are produced by the overtones.

The positions of the nodes and loops in vibrating columns of air can be investigated by means of an arrangement devised by Koenig, and called a manometric capsule or flame.

A hole is made in the side of the tube AB (Fig. 263), and over this hole is stretched a thin india-rubber membrane C. A small metal, or wooden, capsule D covers the membrane, leaving a small enclosed space G. Ordinary coal gas is supplied through the tube E, and escapes through F, where it is lighted. If the pressure within the pipe alters, the membrane C will be forced in and out, causing the pressure in G to vary also. The results of the variation of the pressure of the gas in G will be to cause the size of the flame to vary, when the pressure in G is increased the size of the flame increases, while when the pressure in G decreases, so also does the size of the flame. Since the variations in the pressure inside the sounding-pipe occur with great rapidity, the changes in size of the flame cannot be observed when the flame is looked at directly, on account of the persistence of vision, for the images of the small and large flames made on the retinæ overlap. In order to overcome

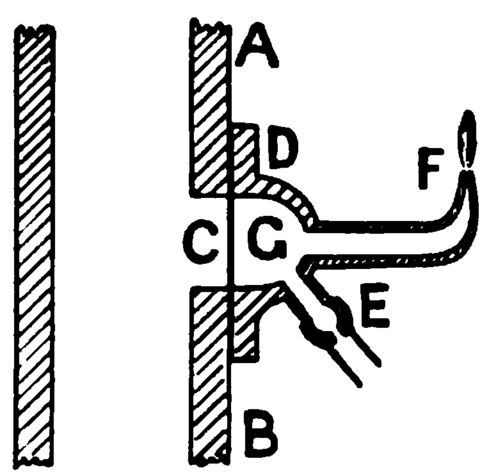


FIG. 263.

this difficulty, the flame, instead of being observed directly, is looked at by reflection in a mirror, shown in Fig. 264, which can be rotated about a vertical axis, so that the images of the flame are no longer superposed. If the hole in the side of the pipe coincides with a loop, then there will be no variations in pressure, and hence the manometric flame will not vary in size, and when viewed in the rotatory mirror will show as a continuous band of light. If, however, the hole is at a node, the flame, when viewed in the rotatory mirror, will be broken up into a serrated appearance, as shown in Fig. 265.

FIG. 264.

We have in the above discussion supposed that a loop was formed exactly at the open end of a pipe. This, however, is not accurate, for it is only at a little distance beyond the end of the pipe that no changes in density occur. If there is a flange at the open end of the pipe,

FIG. 265.
(From Ganot's "Physics.")

as shown in Fig. 266, the loop occurs at a distance of 0.82 R outside the end of the pipe, where R is the radius of the pipe. If there is no flange, the loop is at a distance of 0.57 R from the end. Hence the distance between the open end of a pipe and the nearest node is always less than half the distance between any two consecutive nodes, or less than $\lambda/4$, where λ is the wave-length of the note given by the pipe. The effect of this correction for the open end is virtually to lengthen the pipe, but this will not alter the relative pitches of the overtones. Since, however, the correction for an open pipe will have to be applied at both ends, its virtual length will be $l + 2a$, where a is the correction for the end, and the wave-length of the note

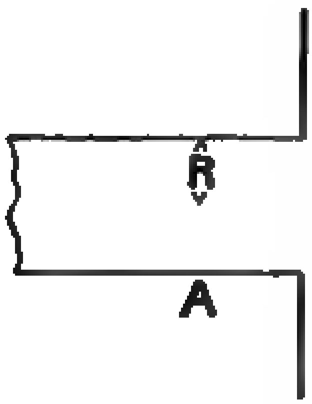


FIG. 266.

emitted will be $2(l+2a)$, while the virtual length of a closed pipe of length l will be $l+a$, and the wave-length of the note emitted will be $4(l+a)$. Hence the interval between the notes given by an open and a closed pipe of the same length (l) will be

$$\frac{4(l+a)}{2(l+2a)},$$

and this is less than 2. Hence the open pipe, instead of giving the octave of the note given by the closed pipe, as the elementary discussion previously given would lead us to expect, gives a note somewhat *lower* than the octave.

307. Organ-Pipes.—The most familiar case of the vibration of columns of air occurs in the case of organ-pipes. An organ-pipe consists of two parts: (1) a tube enclosing a column of air which is set in vibration, and which governs the pitch of the note emitted; and (2) an arrangement for setting this column of air into vibration and maintaining the vibrations when started. There are two distinct ways in which the vibrations of the air column can be started and maintained. In one of these air is forced through the channel A, Fig. 267 (*a*), and the stream of air strikes against the bevelled lip B of the pipe. The stream of air striking this edge sets up vibrations in the air contained within the body of the pipe, in the same way that vibrations can be set up in the air contained in the barrel of a key by blowing across the top.

In the other method the air is set in vibration by means of the transverse vibration of a thin plate of metal, C, Fig. 267 (*b*), called a reed, which is fixed at one end, and nearly fills the aperture leading from a box, D, to the pipe E. If air is forced into D it will, in escaping, set the plate C in vibration, and the reed in its motion alternately closes and opens the passage from the box D. It is the impulses derived from this intermittent supply of air which sets the column of air in the pipe into vibration.

(*a*) (*b*)

Fig. 267.

Open organ-pipes are tuned by reducing the size of the open end by bending a sheet of metal so that it covers the opening more or less. The smaller the opening, the lower is the note given. Closed pipes are tuned by forcing in more or less the plug which constitutes the closed end, and thus altering the length of the pipe.

If the pipe is not very narrow, the note given when it is blown gently is very nearly a pure tone. If, however, the pipe is narrow or the wind pressure is great, the pipe will give a note in which the first overtone is very marked. When very strongly blown, the first and second overtones

are so strong as to completely drown the fundamental, which is very weak.

The clarionet consists essentially of an open reed-pipe, the length of which can be altered by opening and shutting various valves. The note given is very rich in overtones, and it is owing to the presence of these overtones that this instrument owes its characteristic sound.

In the cornet and horn there is no reed attached to the instrument, but the lips of the performer vibrate and perform the functions of a reed. In the cornet, the length of the tube can be altered at will, and thus the different notes can be obtained. In the horn and bugle the length of the tube remains constant, but the performer alters the manner in which his lips vibrate so as to make the pipe give its different overtones, so that in these instruments all the possible notes that can be obtained are the overtones of the fundamental note which the pipe will give.

CHAPTER VI

SUPPLY OF ENERGY TO A SOUNDING BODY—RESONANCE

308. Vibrations Maintained by Heat.—In the case of organ-pipes, which are the only sources of sound which we have considered which are capable of giving a maintained note, the energy necessary to maintain the vibrations, and make up for the energy which is radiated as sound-waves, is supplied by the blast of air used to make the pipe "speak." We have now to consider other methods by which the energy which is communicated to the air as sound-waves by a sounding body can be supplied. In the case of a tuning-fork, say, the energy necessary to supply the sound-waves is derived from the loss of energy of motion of the prongs, so that the sound gradually dies out. In some cases, however, the necessary energy is supplied in the form of heat. The most familiar case of sound being produced by heat is Trevelyan's rocker. This instrument consists of a piece of copper or iron, the cross section of which is shown at CD (Fig. 268), which is heated and then rested on a block of lead, AB. Under these circumstances the rocker gives out a musical note. The cause of the vibrations is the expansion of the lead owing to the heat conducted from the rocker. Suppose the rocker to be resting on the edge C more heavily than on D, the result will be that heat will flow more rapidly into the lead at C than at D. This heat will cause the lead to expand immediately under the edge C, and this expansion will tilt the rocker over on to the other edge. The conduction of heat will now be greater at the edge D, so that the lead will now expand under it and cause the rocker to tilt back on to the first edge, when the whole process will be repeated. Thus the rocker is set in vibration and gives out a note, the energy being supplied by the heat of the rocker; in fact the arrangement forms a heat engine in which the rocker is the source and the lead block the sink, and some of the heat of the source is converted into energy of motion, while at the same time a portion of the heat passes from a higher to a lower temperature, that is, passes to the sink.

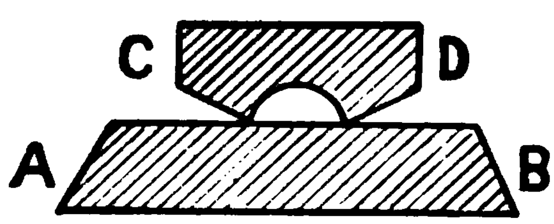


FIG. 268.

If a jet of hydrogen gas be placed within a vertical tube open at both ends, then in general a loud note will be produced, which will continue as

long as the gas jet remains alight. The same phenomenon is exhibited by burning jets of other combustible gases, but to a less marked degree. If the flame is observed by means of a rotating mirror, similar to that used in connection with manometric flames, it will be seen that the flame is in vibration. By using the stroboscopic method of observing the flame, Töpler was able to show that in many cases at one time during each vibration the flame retires inside the jet through which the gas is supplied. It is also found that the length of the gas supply-tube bears an important part in the phenomenon. If the supply-tube is lightly plugged with cotton-wool near the jet the gas flame, although it appears just as usual, is incapable of producing vibrations, while the notes which can be obtained with any given flame depend on the length of the supply-tube and on the nature of the gas. These observations indicate that stationary waves are set up in the supply-tube. The effect of these vibrations in the supply-tube is that the emission of the gas, instead of being uniform, is intermittent, so that the size of the flame, and hence also the supply of heat to the air contained in the tube which surrounds the flame, is intermittent. Now when a column of air is in vibration and heat is supplied to the air at the moment of greatest condensation, this supply of heat will increase the force with which the gas tends to expand, *i.e.* to regain its normal condition of pressure. The effect of this will be similar to that produced when a pendulum is struck a blow at the end of its swing tending to drive it back towards its position of rest, namely, it will tend to increase the amplitude of the vibrations. If, on the other hand, the supply of heat takes place when the air is at its greatest rarefaction, this will tend to resist the return of the air to its condition of rest, and will therefore tend to check the vibrations. Just as in the case of the pendulum, if it is struck a blow tending to check its motion as it is passing through its position of rest, the amplitude will decrease. Hence, if the periodic increase in the size of the flame always occurs at the instant when the air, in that portion of the tube near the flame, is, owing to the natural vibrations of the column of air in the tube, at its maximum condensation, the amplitude of the vibrations will be increased or at any rate maintained. If, however, the increase in size of the flame occurs sometimes at the instant of maximum condensation and sometimes at that of maximum rarefaction, that is, if the natural periods of the column of gas in the supply-tube and of the column of air in the tube are not commensurate, the heat will sometimes assist the vibrations and sometimes oppose. Hence, under these circumstances, the vibrations of the air in the tube will not on the whole be maintained by the heat, and so will die out. It will thus be evident why it is necessary that the length of the supply-tube and the position of the flame should bear definite relations to the length of the tube in order that a sound may be produced. When a plug of cotton-wool is placed in the supply-tube vibrations can no longer take place in the gas contained in the tube, and so the *variations* in the

size of the flame, which are necessary if the vibrations in the air column are to be kept up, are not produced.

309*. The Energy of a Vibrating String.—When a string is vibrating transversely, it possesses energy due to its condition. When it is at its maximum elongation on either side of its position of rest, it is momentarily at rest, and so its energy is entirely potential, that is, is stored up owing to the deformation of the string. When the string is passing through its position of rest, its energy is entirely kinetic.

Let the mass of unit length of the string be m, then as the string vibrates each unit of length will vibrate backwards and forwards in a simple harmonic motion. Let the amplitude of the vibrations executed by an element of the string of unit length, and therefore of mass m, be a, and let its displacement from its position of rest at a given instant be x. Then, as shown in § 51, the velocity with which the element is moving is—

$$2\pi n\sqrt{a^2 - x^2},$$

where n is the frequency of the vibrations executed by the string. Hence the kinetic energy of the element is

$$2\pi^2 n^2 m(a^2 - x^2).$$

Now the acceleration with which the element is moving when its displacement is x is

$$4\pi^2 n^2 x,$$

hence the force acting to produce this acceleration is equal to the product of the mass into the acceleration, or

$$4\pi^2 n^2 mx.$$

This is the force of restitution when the displacement is x, and we see that it is proportional to the displacement x. Hence if we draw a line, OP (Fig. 260), to represent the connection between the displacement and the force of restitution, it will be a straight line passing through the origin, for when the displacement is zero, so is the force of restitution.

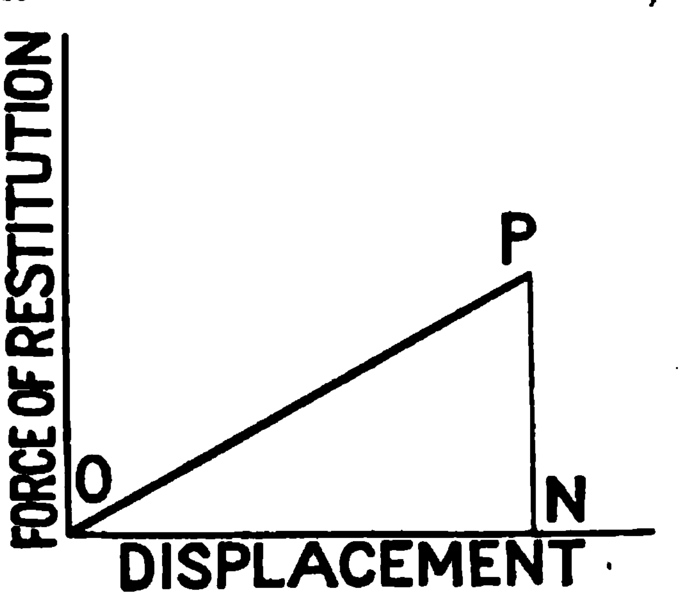


FIG. 269.

If NP represents the force of restitution when the displacement is x, the work which has been done against the force of restitution to displace the element to x is equal to the area of the triangle OPN (§ 77). Hence, as the potential energy when the displacement is x is equal to the work done in displacing the element to x from its position of rest, the potential energy is

$$\tfrac{1}{2}x.\overline{NP}.$$

But $\overline{NP}$ is the force of restitution when the displacement is x, so that the potential energy is

$$2\pi^2n^2mx^2.$$

Hence the total energy, both potential and kinetic, of the element when the displacement is x is given by

$$2\pi^2n^2m(a^2-x^2)+2\pi^2n^2mx^2$$
$$=2\pi^2n^2ma^2.$$

Since this expression for the total energy does not involve the displacement x, we see that the total energy remains constant throughout the vibration, as of course it must, and we simply have changes from the potential to the kinetic form, and *vice versa*, during the motion.

To find the total energy of the whole string we have to add together the energy due to all the elements, so that the total energy is

$$\Sigma 2\pi^2n^2ma^2,$$

where the amplitude a varies from element to element.

To proceed any further we must make some assumption as to the relation between the amplitudes of the different parts of the string. If l is the length of the string and A is the amplitude at the centre, then, if the string is vibrating in its fundamental form, we may represent the amplitude of a point at a distance d from one end by the expression

$$a=A \sin \pi d/l.$$

When $d=o$ or $d=l$, that is, at the ends, a is zero, for sin o and sin π are both zero. When $d=l/2$, that is, at the middle of the string, $a=A$, for $\sin \pi/2=1$. Hence the expression does give us the correct values of the amplitude at the ends and the centre. Substituting this expression for a, we get the total energy equal to

$$\Sigma 2\pi^2n^2mA^2 \sin^2 \pi d/l,$$

or

$$2\pi^2n^2mA^2\Sigma \sin^2 \pi d/l.$$

Now the expression $\sin^2 \pi d/l$ does not involve the amplitude with which the string is vibrating, neither does the expression $2\pi^2n^2m$. Hence the total energy of a vibrating string is *proportional* to the square of the amplitude A with which the centre is vibrating.

Now the only scientific method of measuring the intensity of the vibrations of a body is to consider the energy which the body possesses on account of these vibrations. Hence we see that the intensity of the vibrations of a string are proportional to the square of the amplitude of the vibrations.

By a similar line of argument it can be shown that in the case of all vibrations the energy is proportional to the square of the amplitude. Hence the intensity of all vibrations is proportional to the square of the amplitude.

310. Decrease of the Amplitude of Waves with Increase of Distance from the Source.—Suppose we have a centre of disturbance A within an isotropic medium, so that the wave-fronts are spheres with A as centre. Let R_1 be the radius of one of the spherical wave-fronts, and let the amplitude of the waves as they cross the surface of this sphere be A_1. Similarly let A_2 be the amplitude of the waves when they reach a sphere of radius R_2. Now if we consider a *thin* shell of radius R_1 and thickness x, the energy due to the waves contained in this shell is proportional to the volume of the shell and to the square of the amplitude. Thus the energy is equal to

$$4\pi R_1^2 x . A_1^2 . K,$$

where K is a constant.

Now the waves travel out carrying their energy with them, and when they reach the sphere of radius R_2 the energy contained in a shell of thickness x will be

$$4\pi R_2^2 x . A_2^2 . K.$$

Now the waves which occupy this new shell are the same that some time previously occupied the shell of radius R_1, and so the energy contained within the new shell must be equal to that which was contained within the old. We are here, of course, supposing that the waves can travel through the medium without any of the energy being dissipated as heat, &c., communicated to the medium. Hence, equating the energy contained within the two shells, we get

$$R_1^2 A_1^2 = R_2^2 A_2^2,$$

or

$$A_1/A_2 = R_2/R_1.$$

That is, the amplitude decreases as the distance from the centre of disturbance. The intensity of the wave-motion being proportional to the square of the amplitude, it follows that the intensity decreases as the square of the distance from the centre of disturbance.

This result is, of course, applicable to the case of sound-waves, so that the intensity of a sound varies inversely as the square of the distance from the sounding body. This only applies if the sound-waves are propagated in free air, so that the wave-fronts are spheres. If the sound is propagated along the air contained within a tube, the cross section of the wave-fronts at all distances from the source remains the same, and hence the decrease in amplitude is only due to heat produced by friction of the moving air against the sides of the tube, and such like causes.

311. Damping.—When a vibrating body produces a sound, the energy of the sound-waves which travel out from the body is derived from the energy of vibration of the body; also a certain amount of the energy is converted into heat on account of viscosity of the particles of the body and friction. This loss of energy causes a gradual decrease in the amplitude of the vibrating body, unless energy is supplied to the

body from some external source. If the loss of energy is rapid, so that the amplitude of the vibrations decreases rapidly, the vibrations are said to be *damped.* The extent of the damping is measured by the ratio of the amplitude of one vibration to that of the succeeding one. Thus, if A_1 and A_2 are the amplitudes of successive vibrations, the ratio A_1/A_2 measures the damping. In general this ratio is constant, so that the difference between the logarithms of the amplitudes of successive vibrations is also constant. For if

$$A_1/A_2 = \text{constant},$$
$$\log A_1 - \log A_2 = \log \text{(constant)}.$$

Thus the difference between the logarithms of successive amplitudes measures the damping of the vibrations, and is called the logarithmic decrement.

312. Forced and Free Vibrations.—Suppose we have a pendulum of which the period is one second, and that starting with the pendulum at rest we act upon it with a force which has a period of 1.05 seconds. Let us for simplicity suppose that what happens is that the pendulum is struck a number of small blows, the period of the blows being 1.05. The result of the first blow is to start the pendulum swinging. At the next blow the pendulum will have made a little more than one complete vibration, but the blow will act in the direction in which it is moving, and so will increase the amplitude. At the fifth blow the pendulum will have completed $5\frac{1}{4}$ vibrations, and so will be at its extreme position. At the sixth blow the pendulum will be moving in the opposite direction to the force, and hence the blow tends to check the motion. At the tenth blow the pendulum will simply be brought to rest again. With the succeeding blows the whole operation will be repeated. Suppose now that the period of the blows is gradually reduced. As a result, the number of blows which are struck before they begin to *decrease* the amplitude gets greater, and hence the maximum amplitude attained increases. When the period of the blows is one second, then the pendulum will *always* be moving through its position of rest in the direction in which the force acts at the instant when the blow is struck, and so the blows will always tend to increase the amplitude of the vibrations.

From the above illustration it will be evident why it is that when the period of a periodic force is the same as the natural period of the body on which it acts, then the vibrations set up in this body by the force are very much more energetic than if the period of the force has any other value.

In the above discussion we have supposed that the blows had no effect on the period of the pendulum. This is not, however, quite correct, as unless the blows are struck while the pendulum is passing through its position of rest, the period will be affected. For suppose a blow is

struck the pendulum when it is at its extreme elongation so as to tend to increase the elongation, the effect will be to delay the return swing of the pendulum, and thus increase its period.

From the above discussion, when the force is supposed to be intermittent it will be clear how, when the force is periodic, that is, alternates in direction as well as changes in magnitude, a body may be set in periodic motion by the force. There are two cases to be considered. In the first place, if the natural period in which the body would vibrate, if it were displaced and then left to itself, is the same as the period of the force, the vibrations set up by the force will be very energetic, and the body is said to *resound* to the periodic force. In the second place, if the natural period of the body does not agree with the period of the force, then the vibrations set up in the body are in general of very small amplitude, and they agree in period, not with the natural period of the body, but with the period of the force. Such vibrations, which do not agree in period with the natural period of the vibrating body, but are produced by the action of a periodic force, are called *forced* vibrations.

If when a body is executing forced vibrations the periodic force is stopped, then the body will continue to vibrate, but the vibrations will be of the same period as the natural period of the body, and are said to be free.

If the point of support, A (Fig. 270), of a pendulum is given a periodic to-and-fro motion, we shall have a case of a periodic force being applied to the pendulum. If the period of the to-and-fro motion is the same as the natural period of the pendulum, this latter will be set into violent oscillation. If, however, the period of the to-and-fro motion is not the same as the natural period of the pendulum, the oscillations produced will be of very much smaller amplitude. After a time the pendulum will take up a periodic motion of the same frequency as the to-and-fro motion, the amplitude remaining constant. In this case the motion of the pendulum is forced. If the to-and-fro motion from A′ to A″ is slower than the natural period of the pendulum, the motion of the pendulum will be as shown at (*a*), Fig. 270. The pendulum behaves as if the point of support were at the point P, which is so situated that a pendulum of length PB would have the same period as the to-and-fro motion. If the natural period of the pendulum is less than that of the to-and-fro

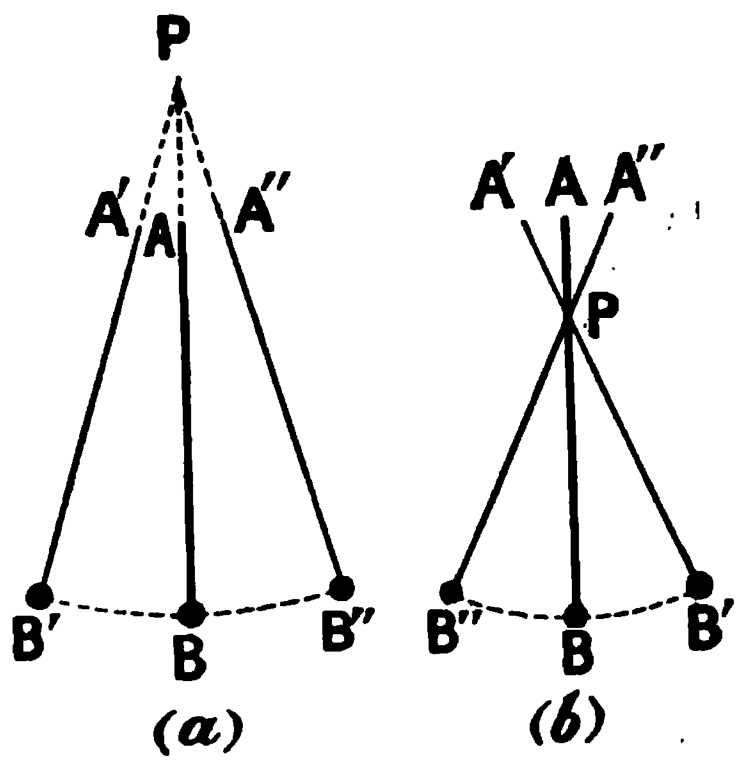


FIG. 270.

motion, the motion of the pendulum is as shown at (*b*), where a pendulum of length $\overline{PB}$ would vibrate in the same period as the to-and-fro motion. If the to-and-fro motion of the point of support is stopped, the pendulum will continue to vibrate, but now with its natural period. This example illustrates very clearly how a body will adjust its period to agree with the period of an impressed periodic force.

A case of forced vibrations occurs if two clocks, which are nearly keeping the same time when they are placed on separate stands, are placed on the same stand. The vibrations of the pendulums then cause the stand to oscillate slightly, so that one pendulum produces a periodic force acting on the other, and, as a result, the clocks, which when at a distance do not keep time, when together keep time exactly. The pendulum of the clock which loses exerts a periodic force, which produced forced vibrations in the faster pendulum, causing it to vibrate slower, while the faster pendulum in the same way causes the slower pendulum to vibrate faster. Thus finally the two pendulums vibrate in periods which are exactly the same, and the vibrations in each case are forced, though they differ in period only very slightly from the natural period. The increase in the amplitude of the vibrations produced by a periodic force, when its period agrees with the natural period of the body, is also very clearly shown by the oscillations of a suspension bridge. Standing in the centre of such a bridge, it will be found that when the steps of a horse have a certain period the bridge is set into violent oscillation, while for all other periods, even when the horse is much heavier, the oscillations are almost imperceptible.

If the free vibrations which a body is able to execute are heavily damped, the amplitude of the vibrations produced by a periodic force will not increase in such a marked manner when the period of the force agrees with the natural period of the body. The reason for this is, that, owing to the damping, the vibrating body loses during each vibration nearly all the energy communicated to it by the corresponding impulse of the force, so that when the force acts next time the body does not possess any great quantity of the energy communicated at the preceding impulse. Thus the energy supplied by the force is not stored up in the vibrating body by the amplitude of the vibrations increasing, but is continually dissipated. An idea of the conditions can be obtained by imagining two pendulums, the bob of one being a ball of paper, and of the other a ball of lead. The paper pendulum is strongly damped on account of the friction of the air, and if we strike it periodic blows it will vibrate, the amplitude of the vibrations being very nearly the same after the first blow as after a number of blows. In the case of the pendulum with the heavy bob it is otherwise, for by suitably timing the blows the amplitude can be gradually increased to a considerable amount.

313. Resonators.—Resonance, that is, the production of vibrations in a body by the action of a periodic force which has the same period as

the natural period of the body, occurs frequently in sound, and plays an important part in music.

The periodic force is produced by the vibrations of one body, and is communicated to the second body, which is called the resonator, by the sound-waves set up in the air, or other medium (solid or fluid) which separates the two bodies. Thus suppose we have two tuning-forks of the same pitch, and that one of these is sounded when held near the other. The sound-waves from the sounding fork will strike the other fork, and, since the frequency of these waves is equal to the frequency of the fork by which they are produced, by their impact they will set the fork in vibration. In this case, since the damping of the vibrations of a tuning-fork is very small, it is necessary for resonance that the pitches of the two forks should be very accurately adjusted to equality. In fact if, when the adjustment has been made so that strong resonance takes place, one fork be warmed, thus causing its pitch to fall slightly, the resonance is almost completely destroyed.

The chief function of resonators in acoustics is to strengthen the amount of the sound, of the particular pitch to which they respond, which is radiated through the surrounding air, that is, to increase the amplitude of the waves produced in the external air by means of which the sound is heard. As an example of this action of a resonator, we may take the case of the increase in the loudness of the note given by a tuning-fork by means of a resonator. From the fact that the prongs of a fork have not a very great area, they are not capable of setting any great quantity of the surrounding air in violent vibration, for the air on the side towards which the prong is moving can slip round the edge of the prong, and thus partly fills up the rarefaction which is being produced on the other side of the prong. In addition, the interference which takes place between the waves emitted from the two prongs (§ 294) reduces the intensity of the motion produced in the surrounding air. If, however, the sounding fork is held near the open end of a closed pipe, of which the natural period is equal to that of the fork, that is, its length is equal to $\frac{1}{4}$ of the wave-length of the note given by the fork, then this pipe will be in resonance with the fork, and the column of air within the pipe will be set into vibration. Now if the open end of the pipe is not very small, the vibrations of the air inside will, at the open end, set the external air into vibration much more powerfully than the fork alone did. The result is that on bringing such a resonator near a sounding fork, the intensity of the sound heard is very much increased. The same kind of effect can be easily noticed in the case of the vibrations of a string. Here, again, the surface of the vibrating body is small, so that it is incapable of setting any great mass of air into vibration, and in addition the waves produced on the two sides of the string are in opposite phase, for when a condensation is being produced on one side of the string a rarefaction is being produced on the other side, and these two sets of waves interfere.

By holding a pipe near, of which the natural period agrees with that of the string, or, what is better, connecting one end of the string to the walls of the pipe, so that the vibrations of the string are communicated to the walls, and by them to the air contained within the pipe, the column of air is set into vibration, and communicates its vibrations to the external air much more powerfully than the string alone is capable of doing.

The above examples of the strengthening of the sound produced by a vibrating body, which are cases of true resonance, must not be confounded with others where the vibrating body is able to set up *forced* vibrations in a body, and the increase in the loudness is due to these forced vibrations. Thus when the stem of a vibrating tuning-fork is pressed against the wooden top of a table, the loudness of the sound produced is greatly increased. In this case the fork sets the table into forced vibrations which have the same frequency as the fork, and the table, on account of its large surface, is able to produce violent sound-waves in the air. That this is not a case of resonance is shown by the fact that the table acts equally well in the case of forks of all frequencies.

Since the resonator owes the energy necessary to set it into vibration to the sounding body, and since the increased loudness of the sound emitted when the resonator is present means that more energy is being given out to the external air, it follows that the sounding body must lose its energy of motion more rapidly when a resonator is present than it does when no resonator is present. That this is so can be easily shown by sounding a tuning-fork first without the resonance-box belonging to it, and then with the box, when it will be found that the vibrations of the fork continue for a much longer time without the resonator than they do with it.

Another important use of resonators is in assisting the ear to resolve a compound note into its simple tones, for if the note contains a tone which is the natural tone of a resonator, the resonator will be set into vibration by this tone ; while if this tone is absent from the note, the resonator will not respond. We shall see in a subsequent section how this analysing property of resonators may be utilised.

A resonator, just as any other body, can be set in vibration, not only in its fundamental mode, but also so as to produce overtones. Hence a resonator will resound to a tone having the frequency of any of its overtones. Now one of the uses of resonators is in the production of pure tones, that is, a resonator is used to strengthen the fundamental tone of a note without strengthening the upper partials which may be present at the same time. If, however, the upper partials of the natural vibrations of a resonator were of the same pitch as the upper partials of the note, these would be strengthened by the resonator as well as the fundamental. Thus, in choosing a resonator, it is important that while

it resounds to the fundamental it should not resound to the upper partials. We have seen that in the case of a closed cylindrical pipe the frequencies of the upper partials are 3, 5, 7, &c., times the frequency of the fundamental, so that if the sounding body, say a fork, has partials of any of these frequencies, these partials will be strengthened. In the case of resonators of the shape shown in Fig. 274, the frequency of the overtones is very high compared to the frequency of the fundamental, and so, if, as is the case with a fork, the upper partials which are present in any strength are comparatively low, the resonator will not be able to strengthen them. Thus for producing a pure tone the resonators of the form shown in Fig. 274, rather than simple closed pipes, are to be preferred. The periods of the fundamental tones of such resonators cannot usually be calculated, but the following general considerations apply: For a given opening or mouthpiece the pitch is mainly dependent on the volume of the enclosed air, while in the case of resonators without necks the influence of the mouth depends on its area.

314. Kundt's Experiment.—If a rod AB (Fig. 271) is clamped at its middle so that one end projects into a tube CD, the end of the rod being fitted with a light piston which fits loosely into the tube, then, on causing the rod to vibrate longitudinally, this piston will vibrate backwards and forwards and will set up vibrations in the air contained in the tube CD.

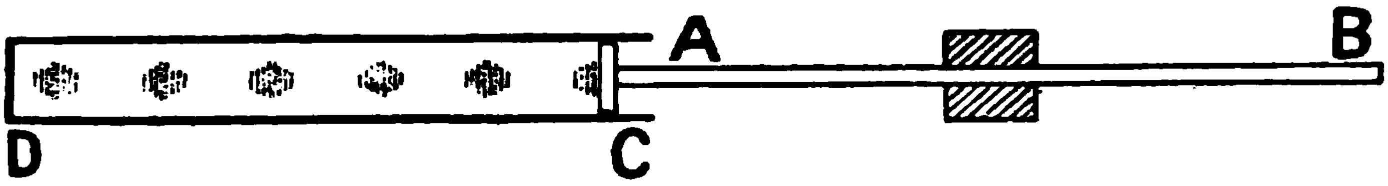


FIG. 271.

The waves in the air in the tube will be reflected from the end D of the tube, and the direct and reflected waves will set up stationary vibrations in the air. If we suppose that the tube is closed at D, this point will be a node, and there will be a series of nodes along the tube at distances equal to $\lambda/2$ from one another, where λ is the wave-length, in the gas which fills the tube, of the tone having a frequency equal to that of the rod. If the position of the end of the rod is adjusted so that the piston B is at a loop of these stationary vibrations, the motion of the piston will have its maximum effect in increasing their amplitude, and they will be so intense that if a light powder, such as lycopodium or cork filings, be strewn inside the tube, it will, by the vibration of the air or other gas, be collected in very characteristic transverse ridges at the loops. The explanation of the formation of these ridges is beyond the scope of this work, so we must content ourselves with referring the reader who wishes to pursue the subject to Rayleigh's "Sound," vol. ii. p. 46. By measuring the distance between consecutive loops, we obtain the value of $\lambda/2$ for the tone produced by the rod in the gas, and this represents the space traversed by a sound-wave in the gas during the time the rod makes

half a vibration. If n is the frequency of the rod, this will also be the frequency of the vibrations in the gas, so that if v is the velocity of sound in the gas, we have $v=n\lambda$, or $v=2nl$ where l is the distance between two of the loops in the tube. If the rod is giving its fundamental, then the wave-length of the sound in the rod is (§ 304) equal to $2L$, where L is the length of the rod. Hence for the material of which the rod is composed we have $V=2nL$, while for the gas in the tube $v=2nl$. Therefore

$$V/v=L/l.$$

Thus, by measuring the ratio of the length of the rod to the distance between two loops, we can calculate the ratio of the velocities of sound in the material of the rod and in the gas. If we know the frequency n of the rod, the velocity of sound in the gas can be calculated, so that by filling the tube with various gases we can obtain the velocity of sound in these gases. Without knowing n we can, by simply comparing the values of the wave-length, obtain the ratio of the velocities in the different gases. For if v and v' are the velocities of sound in two gases, and λ and λ' the wave-lengths corresponding to the tone of frequency n given by the rod, we have $v=n\lambda$, and $v'=n\lambda'$, so that

$$\frac{v}{v'}=\frac{\lambda}{\lambda'}=\frac{l}{l'},$$

where l and l' are the distances between consecutive loops as given by the Kundt's tube.

In the following table the value of the velocity of sound for some gases at a temperature of 0° C. is given. (As has been pointed out in § 286, the velocity is independent of the pressure.)

VELOCITY OF SOUND IN GASES AT 0° C.

Air		330.6	metres	per sec.
Hydrogen		1286	„	„
Oxygen		317	„	„
Carbon dioxide		262	„	„
Coal gas . . .	(about)	490	„	„

CHAPTER VII

AUDITION, COMBINATION TONES, CONSONANCE, AND VOCAL SOUNDS

315. Audition.—In considering the subject of the effects of sounds on the ear, we shall deal exclusively with the physical side of the subject, referring the reader to books on physiology for an account of that aspect of the subject.

The ear is not capable of detecting as sound the presence of air-waves of all frequencies, but it is only when the frequency of such waves falls between certain limits that the ear is able to distinguish their presence, and we experience the sensation we call sound. These limits are, however, neither of them well defined. Helmholtz concluded from his experiments that the lowest frequency which causes the sensation of a *musical tone* is about thirty vibrations per second. In forming any such estimate, it is very difficult to obtain a tone in which we may be quite certain no overtones are present; for, if they are present, what is actually heard may be the overtones and not the fundamental.

The upper limit of audibility is even more uncertain, for not only does it vary very much with the observer, but there is the added difficulty that it is very hard to determine the frequency of notes of very high pitch. The upper limit of audibility for normal ears appears to be somewhere between 10,000 and 20,000 vibrations per second. Estimates of *pitch* cannot, however, be made above a frequency of about 4000.

A closely related subject is the amplitude of the sound-waves in air necessary for audition. As a result of some experiments on the distance to which a whistle could be heard when a measured power (§ 78) was employed in maintaining the sound, Rayleigh came to the conclusion that under favourable circumstances the ear is able to detect a sound, if the amplitude of the sound-wave exceeds 10^{-8} cm.[1]

The direction from which a sound comes can be judged with considerable accuracy, and although the exact method by which we are able to make this estimate of direction is not known, there is no doubt that we are very largely guided by the effect of the sound on the two ears; probably the slight difference of the intensity with which the sound reaches the two ears is at the base of all such judgments.

[1] Intermittent sounds can be detected by the ear when a continuous sound of the same amplitude is inaudible.

316. Beats.—We have seen in § 54 that when we combine two S.H.M.'s of very nearly the same frequency, of which the displacements take place in the same direction, the resultant motion is periodic, and the amplitude of the motion undergoes periodic variations, caused by the displacement due to the two motions sometimes being in the same direction, and thus aiding each other, and at other times being in opposite directions, and so opposing each other. In the case of sound, we may obtain a similar result, for when two tones, whose frequencies do not differ by more than about sixteen vibrations per second, are sounded together, a periodic waxing and waning of the sound due to the two tones occurs. Under such circumstances the tones are said to *beat*. The production of beats may be illustrated objectively by an arrangement similar to that used to produce Lissajous' figures (§ 301). The two forks are arranged to vibrate in the same plane, so that the amplitude of the movement of the spot of light on the screen is equal to the sum of the amplitudes due to each fork separately. It will then be seen that at each beat the spot of light is drawn out into a line, while

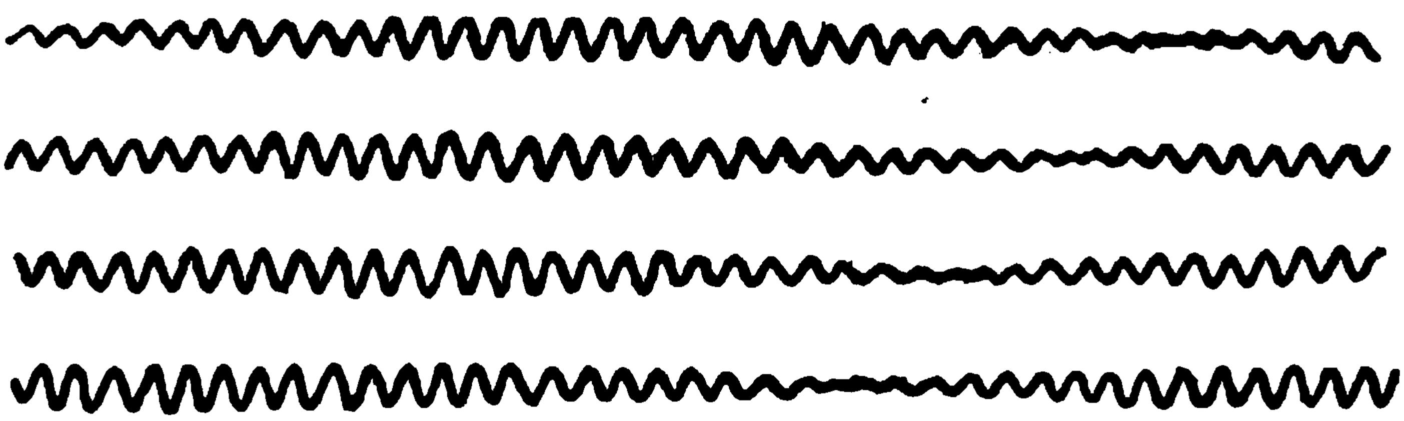


FIG. 272.

half-way between the beats the spot appears round. By projecting the spot of light on a photographic plate, which is moved in a direction at right angles to the plane of vibration of the forks, a curve such as that given in Fig. 272 is obtained, in which the effect of the beats is very clearly shown.

Let one tone (A) make x vibrations, while the other (B) makes $x+1$. If then we start when the two are in the same phase, the phase of the tone B will gain on that of A, till, at the end of x vibrations of A, B will have made $x+1$ vibrations, and so they will again be in the same phase, and the sound will be a maximum. Let the frequency of the tone A be nx, and hence that of B $n(x+1)$. Now from one maximum of sound to the next A makes x vibrations, so that the number of maxima in a second will be n, or there will be n beats per second. But the difference between the frequencies of A and B is $n(x+1)-nx$ or n, so that the number of beats per second is equal to the difference in the frequencies of the two tones.

Starting with two tones in unison, and increasing the frequency of one of them, beats will be produced which are at first slow, but increase

in frequency as the difference in the frequencies of the tones increases. After a time the frequency of the beats will be so great that the ear ceases to hear them as such, and the only sensation is one of discord. While the beats are still audible, the ear is unable to distinguish the separate tones which are producing the beats, but when the beats cease to be distinguishable, then the ear can detect the existence of the two separate tones.

Koenig has advanced the theory that in the case of the beats produced by pure tones, such as those given by massive tuning-forks when lightly bowed, there are really two sets of beats. Thus suppose we have a fork giving 64 vibrations per second, this, according to what we have said above, will give 8 beats per second with a fork of which the frequency is 72. According to Koenig it is, however, possible to obtain 8 beats per second, if the second fork has a frequency of 120. In the first case we have $72-64=8$, and in the second case $64\times2-120=8$, so that beats occur not only where the frequencies of two notes are nearly the same, but also when the frequency of the higher note is nearly equal to the frequency of the octave of the fundamental. Of course the presence of this superior series of beats is explained, if we suppose that the octave of the lower tone is given by the tuning-fork as an overtone, so that the beats heard are really due to the combination of this overtone with the higher fork. Koenig, however, maintains that the tones he uses are pure, that is, that no overtones are present.

Koenig's two series of beats are related to the primary tones in the following manner. The frequency m of the higher tone must lie between two multiples, a and b, say, of the frequency n of the lower tone, where an is less than m, and bn is greater than m. Then the two series of beats which may be produced are an inferior series, in which there are $m-an$ beats per second, and a superior series, in which there are $bn-m$ beats per second. Thus if the two tones had frequencies of 40 and 74, then $a=1$ and $b=2$, since 74 lies between 40×1 and 40×2. Hence the possible beats will be an inferior beat, of which the frequency is $74-40$ or 34, and a superior beat, of which the frequency is $80-74$ or 6. Both sets of beats are not, however, usually audible at the same time.

317. Combination Tones. — There are other phenomena besides the beats or throbbing sensation, which are due to the simultaneous production of two tones. For under certain circumstances, when two tones are sounded simultaneously, the ear is able to detect, in addition to the two primary tones, other musical tones, which are due to the combined effect of the two primary tones. These additional tones are called *combination tones*. There are three kinds of combination tones, namely, difference tones, summation tones, and beat tones.

The difference tone is a combination tone the frequency of which is equal to the difference in the frequencies of the two primary tones.

Thus if the frequencies of the primary tones are m and n, the frequency of the difference tone is $n-m$. In addition to this difference tone, which is called the difference tone of the first order, this combination tone can itself form difference tones with either of the primary tones, thus giving rise to difference tones of the second order, and so on.

The summation tones are less easily obtained than the above, and have a frequency equal to $n+m$. Summation tones of a higher order than the first can also be obtained by the combination of the first order summation tone with one of the primary tones, &c.

When considering the result of the superposition of the effects of two systems of waves in air or any other medium, we have assumed that the displacement of any particle due to the two systems is so small that the force tending to restore the particle to its position of rest is accurately proportional to the displacement. If, however, the wave-systems are so energetic, and the displacement of the particles of the medium so large that this proportionality no longer holds, then von Helmholtz has shown that, in addition to the two primary wave-systems of frequency m and n, there will be produced two secondary wave-systems of which the frequency will be $n-m$ and $n+m$. These secondary systems will correspond to the difference tone and summation tone respectively.

The condition which von Helmholtz presupposes is that the two primary tones cause the *same* body or the *same* portion of air to vibrate very violently. In the case when the two tones are produced by a syren or by two harmonium-reeds, the vibrations produced in the air are apparently sufficiently violent to cause the formation of the combination tones, for von Helmholtz and also Prof. Rücker have shown that when the two primary tones are produced by these sources, the summation and difference tones are capable of causing a suitably tuned resonator to "speak." In other cases the body in which these tones are produced may be the ear itself, for the series of small bones and membranes which convey the sound from the outside drum to the nerve terminations form an arrangement such that, when violently disturbed, the restoring force would no longer be accurately proportional to the displacement.

Koenig has shown that when the beats produced by two tones are sufficiently rapid they combine to form a combination tone, which he calls a beat tone. There will be two series of these beat tones, one corresponding to the inferior series of beats of frequency $n-m$, and the other belonging to the superior series having a frequency $bm-n$. It will be seen that the inferior beat tone corresponds in frequency to the difference tone, and Koenig considers that the difference tone is really a beat tone, *i.e.* is produced by the beats having become so rapid as to form a tone. He also maintains that if the primary tones are really pure the summation tone is never heard, and that when it is heard it is due to beat tones produced between some of the upper partials of the primary notes.

The whole subject of combination tones is one of considerable difficulty, and we must content ourselves with the above very meagre statement of some of the more prominent facts.

318. Consonance and Dissonance.—When dealing in § 289 with the musical scale, we referred to the fact that some intervals are consonant, *i.e.* produce an agreeable effect when the two tones are sounded together, while others are dissonant, or produce an unpleasant effect on the ear. We are now in a position to consider this question rather more in detail.

The end organs of the auditory nerve are situated on a fine membrane, called the basilar membrane, which is situated in the internal bony chamber of the inner ear (the cochlea). According to von Helmholtz's theory of audition, the different threads of the basilar membrane act as resonators, each thread responding to vibrations of a definite frequency, and that as a consequence of the vibration of these threads the corresponding nerve terminations are stimulated, and produce in the brain the sensation which we call sound. Thus when the disturbance produced by the sounding of a musical note is conveyed to the liquid which fills the cochlea, the threads which have a natural period agreeing with the period of any simple tones which are present in the note will be set into vibration. If the sound consists of a simple tone, then only one, or at any rate a very small proportion, of the threads are set into vibration, and the brain is not able to resolve the sound into any simpler sensation.

When two tones of nearly the same pitch are sounded simultaneously, then some of the fibres will respond to both tones, but the vibrations set up will, on account of the production of beats, be intermittent in character. If this intermittence is sufficiently slow, the beats will be heard as separate maxima of sound. If, however, the beats are rapid, so that the fibres have not time to completely come to rest, or at any rate if there is not time for the nerve to recover completely between the stimuli, the effect will be noticed as a roughness or discord. When, however, the interval between the tones is so great that none of the fibres are set in vibration by *both* tones, then the sense of roughness will vanish.

If we further suppose that each thread may be capable of vibrating in more than one way, say the overtones are the harmonics of the fundamental, so that any given fibre would respond to a tone, and its octave, twelfth, &c., we can understand how it is that the interval of the octave is so consonant, and it would further explain why a tone and another, the frequency of which is nearly the same as the octave of the first, produce discord.

Starting with two tones of the same frequency, and then gradually increasing the frequency of one of them, the number of beats produced gradually increases. Very slow beats are not very unpleasant, but as the frequency of the beats increases so does the unpleasantness, till for a certain number of beats per second the discord is a maximum. If the

number of beats is still further increased, the unpleasant sensation gradually decreases. The phenomenon is similar to that which occurs in the case of vision. If the intensity of a light alters, that is, the light flickers, the unpleasant sensation produced first increases as the frequency of the flickers increases, reaches a maximum and then decreases, and if the frequency is more than twenty per second all sense of flicker is lost, and the illumination appears continuous. The frequencies of the beats for which the discord is a maximum, and for which the sensation becomes continuous, vary with the absolute frequencies of the tones which give the beats. This can be clearly perceived by sounding the following intervals on the piano, in which the number of beats per second produced is the same, but while the first interval is very discordant, the last is quite concordant.

Interval.	Notes.	Frequencies.	Number of Beats per Second.
Semitone	*b′, c″*	495, 528	33
Tone	*d′, e′*	297, 330	33
Minor third	*e, g*	165, 198	33
Major third	*c, e*	132, 165	33
Fourth	G, *c*	99, 132	33
Fifth	C, G	66, 99	33

Mayer, who has examined this subject, gives the following series of values for the frequencies of the beats when discord is a maximum and when the roughness vanishes respectively :—

Frequency of Lower Tone.	Number of Beats per Second.	
	When Discord is a Maximum.	When Roughness Vanishes.
64	6.4	16
128	10.4	26
256	18.8	47
384	24.0	60
512	31.2	78
640	36.0	90
768	43.6	109
1024	54.0	135

When considering the accord or discord produced by compound tones, such as ordinarily occur in music, the presence of the upper partials must be taken into account, for although the fundamentals may be in accord, the upper partials of the notes may produce discord.

Of course the effects of the upper partials, in giving dissonance for any given musical interval, will depend on the strength of the partials

and on the relation which their pitch bears to that of the fundamental, so that the source of two musical notes has to be taken into account when considering consonance. For simplicity we shall, however, suppose that the notes are such that the partials consist of the first seven harmonics of the fundamental. Under these circumstances the frequencies of the fundamental and the overtones are represented by the numbers 1, 2, 3, 4, 5, 6, 7, 8.

In each case we will assume the tonic to have a frequency of 256, and will then examine the frequencies of the partials of this tonic and of notes which together with it produce some of the commoner musical intervals. In the first place, if the interval is the octave, so that the frequencies of the two fundamentals are 256 and 512, the frequencies of the partials are shown in the following table :—

	Tonic.	Octave.
Fundamental	256	512
1 overtone	512	1024
2 "	768	1536
3 "	1024	2048
4 "	1280	2560
5 "	1536	...
6 "	1792	...
7 "	2048	...
8 "	2304	...

A consideration of this table will show that the smallest difference in frequency (except unison) between two partials of these two notes is 256. Hence, since for discord this difference must be about 19, we see that not only the fundamentals, but also the upper partials, are consonant. The effect of the second note is to strengthen the even partials of the lower note.

The following table contains the frequencies of the partials of the higher note for some other intervals, the tonic having, as before, a frequency of 256 :—

	Tonic.	Fifth.	Fourth.	Major Sixth.	Major Third.	Minor Third.
Fundamental	256	384	341	427	320	307
1 overtone	512	768	682	854	640	614
2 "	768	1152	1023	1281	960	921
3 "	1024	1536	1364	1708	1280	1228
4 "	1280	1920	1705	2135	1600	1535
5 "	1536	2304	2046	2562	1920	1842
6 "	1792	2688	2387	...	2240	2149
7 "	2048	...	...	...	2560	2456
8 "	2304	...	...	...	...	...

In the case of the fifth, it will be observed that the difference in frequency between the fundamentals is 128, while this number also expresses the smallest difference which occurs between any of the particles.

In the fourth, the smallest difference in frequency is 83. There is, apparently, a difference of 2 between the seventh overtone of the fundamental and the fifth overtone of the higher note, but this is because for simplicity we have taken the frequency of this note as 341, when it ought to be 341.3.

In the major sixth, the smallest difference is 84.

In the major third, the smallest difference is 84.

In the minor third, the smallest difference is 50.

Now of the intervals considered above the most consonant is the fifth, and the consonance decreases as we pass to the minor third. This decrease in the consonance is accompanied by a decrease in the smallest difference in frequency of the partials of the two notes, so that in the case of the minor third we are approaching the limit (47 beats per second) at which discord begins.

Next let us take a case where, although the difference between the frequencies of the fundamentals is greater than in several of the cases above, yet the consonance is not so good, and see whether we can account for this dissonance by the production of beats between the partials. For this purpose we may take the notes *g*♭ and *g*♯, and a slightly mistuned fifth.

	c	*g*♭	*g*♯	Mistuned Fifth.
Fundamental	256	369	400	380
1 overtone	512	738	800	760
2 "	768	1107	1200	1140
3 "	1024	1476	1600	1520
4 "	1280	1845	2000	1900
5 "	1536	2214	2400	2280
6 "	1792	...	...	...
7 "	2048	...	...	...
8 "	2304	...	...	...

In the case of *c* and *g*♭, the difference between the fundamentals is 113, and so these tones will not produce discord. The second overtone of *c* and the first overtone of *g*♭, however, differ by 30, and are therefore dissonant, and it is to the beats produced by these that the dissonance of the interval is due. In the case of *c* and *g*♯, the second overtone of *c* and the first overtone of *g*♯ differ in frequency by 32, while the sixth overtone of *c* and the fourth overtone of *g*♯ differ by 48, and the dissonance of the interval is thus accounted for. In the untrue fifth there will be eight beats per second between the second overtone of the lower note and the first overtone of the higher, sixteen per second between the

fifth overtone of the lower and the third overtone of the higher, and twenty-four between the eighth overtone of the lower and the fifth of the higher. Hence it will be seen why it is that an untrue fifth is dissonant, and how the ear is able to detect want of correctness in such an interval.

We may in the same manner examine the other intervals, but this task is left to the reader.

Since, when an interval is untrue, those partials of the two notes which, if the interval were true, ought to be in unison, will be in a condition for producing beats, it follows that the greater the number of common partials, and the stronger these partials are, the greater will be the discord produced by mistuning the interval, and so the greater the accuracy with which the ear can adjust such an interval. In the case of the perfectly consonant interval, the octave, all the partials of the higher note are in unison with partials of the lower. In the fifth, the alternate overtones (1, 3, 5, &c.) of the higher note are in unison with partials of the lower. In the fourth, every second overtone (2, 5, &c.) of the higher are in unison with partials of the lower. In the major sixth, major third, and minor third, one overtone of the higher is in each case in unison with one partial of the lower, but as the consonance decreases, it is a higher and higher, and therefore less important partial that is in unison. Thus an interval is more consonant the greater the number, and more important, that is, the lower, the partials which are common to the two notes.

When, instead of compound tones, simple tones are employed, the above explanation does not account for the fact that while a true octave or fifth is consonant, an untrue, that is, slightly mistuned, octave or fifth is dissonant. It is to be noted that with pure tones the accuracy with which the ear is able to detect an untrue interval is very considerably less than with compound tones, so that the explanation given above is to this extent supported. Helmholtz has explained the dissonance of simple tones as being due to the beats produced by combination tones. Thus, suppose we have an untrue octave, the frequency of the tones being 256 and 515. The number of beats is 259, and therefore cannot produce the discord which is heard. These two tones will produce a difference tone, of which the frequency is $515-256$ or 259, and this tone will give three beats per second with the tone 256. It is to these beats that Helmholtz attributes our power of distinguishing the untrue interval. In the same way, suppose we have an untrue fifth, the frequencies being 256 and 380. The first difference tone has a frequency of $380-256$ or 124, and will not produce a discord with either primary. A secondary difference tone will be produced between the first difference tone and the lower primary of frequency $256-124$ or 132, and this secondary difference tone will produce eight beats per second with the first difference tone, hence the discord.

319. Timbre.—The quality or timbre of the notes given by different instruments is produced by the overtones which accompany the fundamental. In general those notes in which the fundamental is relatively

strong, the overtones being few and feeble, are of a mellow character. On the other hand, notes in which the overtones are numerous and strong are harsher, and have a so-called metallic sound.

In dealing with the effects of the overtones on the timbre of a note, there are three points to be considered: (1) the relative frequencies of the particles, (2) their relative intensities, and (3) the relations that exist between the phases of the partials.

That the first two of these conditions will have an important bearing on timbre is evident from a consideration of the curves given in Fig. 273,

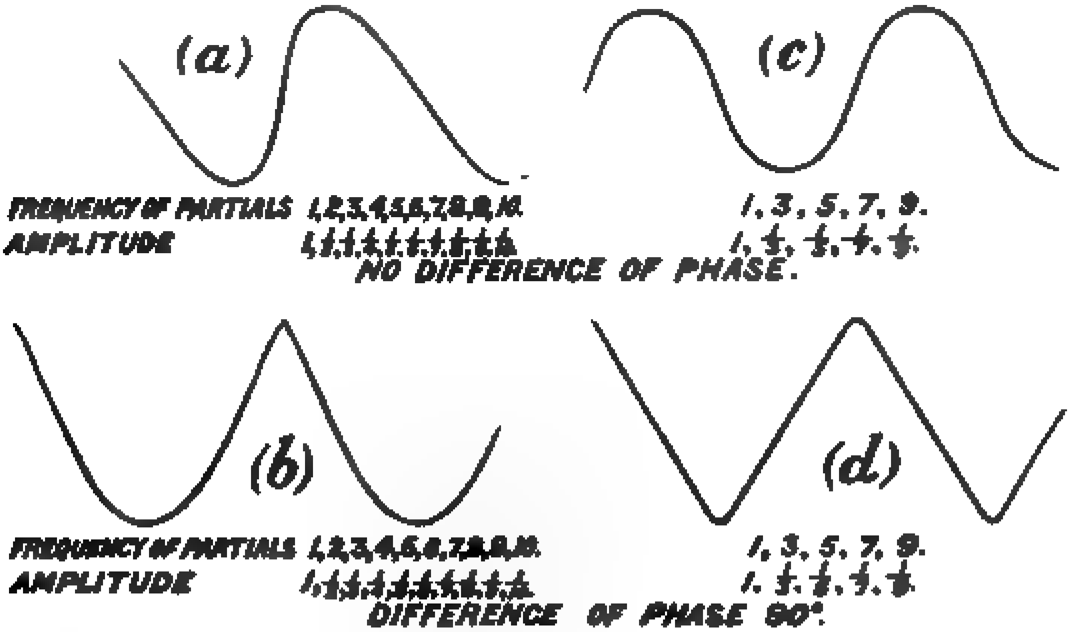


Fig. 273.

(*a*) and (*c*), which represent the resultant curves obtained by the combination of sine curves which are all harmonics of the same curve, but in which the harmonics present are different. In the case of a note in which the corresponding partials are present, we can easily see that the timbre may be very different in the two cases.

In order to determine what partials were present in any given note, Helmholtz used a series of resonators, each of the form shown in Fig. 274. The open end *a* of the resonator is turned towards the source of sound, while the other and smaller opening, *b*, is connected to the ear by means of an india-rubber tube. If the given note contains a partial which corresponds in pitch to the natural pitch of the resonator, then this partial will cause the resonator to "speak."

Fig. 274.
(*From Ganot's "Physics."*)

Koenig has devised the form of resonator shown in Fig. 275, in which the volume of the enclosed air can be altered, and thus the pitch of the note to which the resonator responds also altered.

Instead of using the ear to detect whether a resonator responds to a given note, use may be made of a manometric flame, the capsule of which is connected with the inside of the resonator. On this principle Koenig has constructed the instrument shown in Fig. 276. A series of resonators are tuned so as to respond to a given tone and to its harmonics, each resonator being connected by means of a tube with a manometric capsule. When a note, of which the fundamental agrees in pitch with the pitch of the lowest resonator, is sounded near this instrument all the resonators which agree in pitch with the partials that are present will respond, and the corresponding flames will be affected, which effect can be observed by looking at the flames in a rotating mirror. When an open organ-pipe, of which the fundamental corresponds to the lowest resonator, is sounded gently the fundamental is the only resonator that responds. If, however, the pipe is sounded more strongly, the resonators corresponding to the first five harmonics respond, the response of the third harmonic being stronger than that of the second. If a closed pipe of the same pitch is sounded strongly, then the even harmonics are all absent, as we should expect from the discussion given in § 306. The third harmonic is fairly strong, while the fifth is only feeble.

Fig. 275.
(From Ganot's "Physics.")

In the case of the note given by a bowed violin-string, the first seven harmonics are present, and it is owing to the presence of this large number of partials that the violin owes the piercing character of its notes. In the case of the piano, the partials 1, 2, 3 are fairly strong, while the partials 4, 5, 6 are more feeble, while the position at which the strings are struck is so chosen that the seventh partial is absent. The reason for this is that the seventh partial, when present, forms a dissonance with the sixth and eighth.

The frequency of the sixth, seventh, and eighth partials being $6n$, $7n$, and $8n$, where n is the frequency of the fundamental, the interval between the 6th and 7th is $6/7$, and that between the 7th and 8th is $7/8$. Now neither of these intervals is consonant. If, however, the 7th partial is absent, the interval between the 6th and 8th is $3/4$ (a fourth), and this is consonant. Hence when the 7th partial is wanting, all the partials up to and including the 8th are consonant. Although the 8th and 9th are dissonant, yet since the loudness of a partial decreases rapidly with the order of the partial, the dissonance produced by the 6th and the higher partials is practically negligable.

The partials of an organ-pipe have been investigated by Raps in

another way. Two rays of light are caused to form interference bands (§ 378), and while one ray passes altogether through the external air, the other passes through the air situated at the node of an organ-pipe. The alternate condensations and rarefactions cause the density of the air to alter, and hence the velocity of light passing through the air varies in the same way. The result is that when the pipe is sounded the interference

Fig. 276.
(From Ganot's "Physics.")

bands vibrate backwards and forwards in the same period as the changes of density of the air in the pipe. If then the bands are received on a rotating drum covered with photographic paper, a wavy line will be produced, and from the character of this line the nature of the vibrations of the air in the pipe can be seen. A series of traces obtained in this way from an open pipe blown with gradually increasing wind pressure is

given in Fig. 277. It will be noticed how the first overtone, which is at first absent, gradually increases in intensity as the wind pressure is increased.

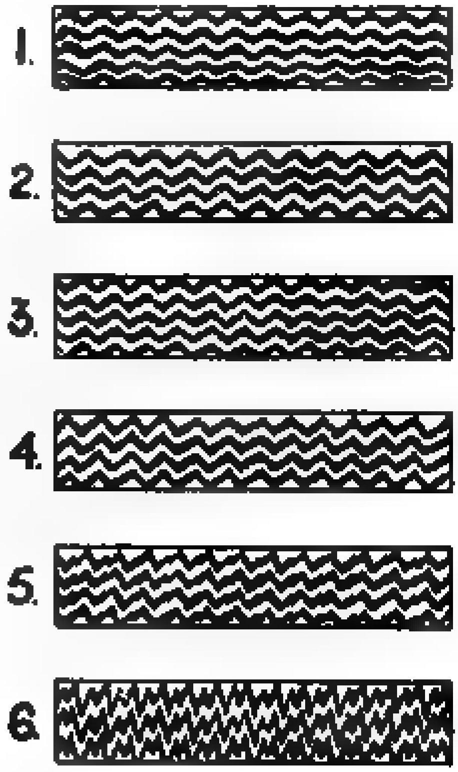


Fig. 277.

Not only did Helmholtz perform the analysis of compound notes into their simple partial tones, but he also performed the inverse operation, namely, building up a note of a given timbre by the combination of a number of simple tones, that is, he performed the synthesis of a given note. The apparatus he used is shown in Fig. 278, and consisted of ten tuning-forks which were tuned so as to give a fundamental of 256 vibrations per second and its first nine harmonics. Each of these forks is arranged in front of a resonator tuned to unison with it. An eleventh fork, K, is maintained in vibration electrically, and is so arranged that it makes and breaks an electric circuit once in each vibration. Each of the other forks is provided with an electro-magnet through which the intermittent current produced by the fork K is sent. The

(From Ganot's "Physics.")

result is that, in the case of fork number 1, the fork is acted upon by

a periodic force which recurs regularly after 1, 2, 3, &c., complete vibrations, and thus keeps the fork in vibration. Fork number 2, in the same way, receives an impulse every other vibration, and so on. Each of the resonators is fitted with a clapper worked by a string attached to a keyboard, by means of which the mouth of the resonator can be closed. When the mouth of the resonator is closed, so that it cannot reinforce the sound of the fork, the sound emitted by the latter is of so feeble intensity, compared to the sound emitted when the resonator is in action, that it may practically be neglected. Thus the compound note heard corresponds to the note produced by the coexistence of the tones given out by the forks, the resonators of which have been uncovered by depressing the corresponding keys. Since the intensity with which a resonator

FIG. 279.

sounds depends on the extent to which the opening has been uncovered, the intensity of the different partials can be altered at will.

Koenig has also performed the synthesis of musical notes, by cutting out the curve obtained as in Fig. 273, by compounding the sine curves corresponding to the different harmonics, taken with their proper relative amplitudes, round the edge of a metal cylinder, such as that shown in Fig. 279. If a jet of air from a narrow slit B be directed on such a toothed wheel, then the passage of the air will vary according to the amount the metal is cut away, and by rotating the wheel a note will be produced. With this instrument, called a wave-syren, Koenig was able to reproduce the characteristic timbre of some musical notes.

The timbre obtained also varies when, the number and amplitude of

the partials being kept the same, the phase relations between these partials are changed in the manner exhibited in Fig. 273. In the curves (*a*) and (*c*) the phases of all the partials are the same, that is, when the fundamental curve is passing through the axis of X from the negative to the positive side, that is, when its phase is 0°, all the curves corresponding to the other partials are also passing through the axis of X from the negative to the positive side, so that their phases are also 0°. In the curves (*b*) and (*d*), when the phase of the fundamental is 0°, that of the partial next above it is 90° behind, so that this curve is at its maximum below the axis of X, while the phase of the next partial is 90° behind this, and so on.

Hence Koenig concludes that, contrary to the results obtained by von Helmholtz, the relative phase of the partials has some influence on the timbre of the note.

320. Production of Vocal Sounds.—The actual organ concerned in the production of the vibrations that constitute a vocal sound are two membranes, called the vocal cords. These membranes are stretched between a series of cartilaginous structures, to which are attached a series of muscles, by means of which the tension of the membranes can be altered. The vocal cords are stretched across the opening of the trachea, which is a tube leading to the lungs, and it is to the vibrations caused in the cords when air is forced between them that vocal sounds are due. The vocal cords in men are thicker than in women and children, so that they vibrate more slowly, and hence produce lower notes.

The sounds produced by the vocal cords are much modified by the effect of the mouth, which acts as a resonator of variable shape and volume. The mouth cavity may, however, not only act the part of a resonator in the sense we have already used the term, but it may also modify the note produced by the vocal cords, even when it is not truly in resonance with any of the lower partials of the note. That such an effect can be produced is shown by an experiment of Koenig's, in which a syren was surmounted by a pipe, and the note produced was examined by a manometric flame, while the ratio between the natural periods of the syren and pipe was altered. It was thus found that the characteristic of the partials present in the resultant note underwent marked changes due to the effect of the resonator, these being very much more marked when the natural period of the resonator coincided with that of one of the partials of the note given by the syren.

Articulate speech is composed of a number of characteristic sounds called vowels, of which there exist almost an infinite number of different kinds, the characteristic being that a vowel sound can be indefinitely sustained without losing its characteristic. In addition to the vowel sounds, there are other sounds called consonants, which are not persistent sounds, being practically only different ways of commencing and ending a vowel sound.

The question as to what it is that gives its character to a vowel sound is a subject about which there has been, and is still, much difference of opinion.

Helmholtz, who investigated this question, found that when the mouth is adjusted for sounding any given vowel, the cavity of the mouth always resounds to the same tone, and that the frequency of the tone depends only on the vowel, and not on the age or sex of the speaker. He was hence led to the conclusion that every vowel sound is characterised by the pitch of the tone (or tones) to which the mouth is adjusted. In this way he found the following characteristic tones for some of the vowels :—

VOWEL CHARACTERISTICS.

Vowel.	Characteristic Tone.	
u as in r*u*de	f	frequency 176
ou	f'	" 352
o as in m*o*re	$b'\flat$	" 476
ah	$b''\flat$	" 932
a as in f*a*ther . . .	d'''	" 1188
e as in th*e*re	b''' and f'	" 1980, 352
i as in mach*i*ne . . .	d'''' and f	" 2376, 176

The difficulty then arises as to how it is that we can recognise the character of a vowel sound when it is sung on very different pitches. This difficulty may in measure be met by supposing that in general one of the upper partials of the note produced by the vocal cords is sufficiently near to the characteristic tone of the vowel to cause the mouth cavity to resound, and it is the strengthening of this partial by resonance which gives the characteristic vowel sound, although the great bulk of the sound produced may be of a different pitch.

By means of a system of resonators, Helmholtz was able to detect the presence of other partials which characterised the vowel sounds, and then, using his apparatus for the synthesis of notes, he was able to reproduce some of the vowel sounds with some degree of success.

More recent observations of the form of the trace in the phonograph, however, seem to indicate that the subject is more complicated than Helmholtz's work would lead us to suppose, and that when a vowel sound is sung on different notes, one of the partials of the note is reinforced, and that the frequency of the partial thus reinforced oscillates between certain limits. Hence the characteristic tone of a vowel is not of perfectly fixed pitch, but changes, within certain limits, in such a way as always to be one of the partials of the note on which the vowel is sung.

321. The Phonograph.—There have been numerous attempts to artificially reproduce human speech, and of these the only one which has

been completely successful is that made by Edison. The latest form of Edison's Phonograph consists of a wax cylinder, on which a small stylus with a sharp cutting edge records the vibrations caused in a thin glass diaphragm by the sound-waves which are incident on the diaphragm. The vibrations of the diaphragm cause the stylus to cut a groove in the wax of variable depth, and this groove forms a permanent record of the vibrations of the diaphragm. In order to reproduce the sounds, the cutting stylus is replaced by a round pointed one, and this stylus is caused to pass along the groove made by the cutting stylus. In this way the rounded stylus is caused to vibrate in exactly the same way as did the recording one, and since it communicates its motion to the glass diaphragm, this latter also is caused to vibrate in the same manner as it did under the influence of the incident sound-waves. The diaphragm communicates its motion to the air, and thus the sounds are reproduced.

The character of the trace made on the wax cylinder has been examined by Hermann, by attaching a mirror to the recording stylus and reflecting a ray of light from this mirror on to a screen. Under these circumstances the movements of the spot of light will show the form of the trace made by the recording stylus on the wax cylinder.

BOOK IV

LIGHT

CHAPTER I

RECTILINEAR PROPAGATION—REFLECTION

322. Scope of the Subject.—The word light is used both in a subjective and an objective sense. Thus when our eye is subjected to certain conditions, we experience a sensation which we call light, while the physical cause of this sensation is also called light. In our examination of the phenomena we shall, however, find that bodies which emit that form of radiation which produces the sensation of light, in general, also emit other forms of radiation which, while physically of the same nature as light, yet do not produce the sensation of light. Thus the flame of a Bunsen burner is almost invisible, still, if we hold our hand on a thermometer near the flame, we find that heat energy is being radiated out all the time. By introducing a piece of fine platinum wire into the flame, the wire will be raised to a white heat, and will produce light radiation. The flame still continues to produce radiant heat, but now, in addition, some radiation is also produced, which can be recognised by our eye. Thus while visible radiant energy will be called light, we shall in this portion of the book also examine the phenomena due to invisible radiant energy, which is physically of the same nature as the light radiation, but for the recognition of which we have no special sense organ.

323. Rays—Geometrical Optics—Physical Optics.—We shall see that in an isotropic medium light is propagated in straight lines, and is due to a wave-motion. A line drawn so as to represent the direction of propagation of the light proceeding from a luminous body is called a ray.[1]

Starting from the observed fact that light travels in straight lines, and the laws of reflection and refraction, it is possible to deduce a number of interesting results which have a direct bearing on optical phenomena by mere mathematical or geometrical methods. The subject of light con-

[1] The word ray is also used to signify "kind of radiation." Thus we speak of heat rays, red rays, blue rays, &c., meaning radiant heat and radiation which produces the sensation of red, blue, &c.

sidered in this way is generally called Geometrical Optics. In geometrical optics no inquiry is made as to the nature or cause of light, neither is any explanation forthcoming of the rectilinear propagation of light. These matters come within the scope of another branch of the subject, called Physical Optics. In the following pages we shall commence by using the methods of geometrical optics. Having in this way obtained a certain familiarity with some of the simpler phenomena, we shall then be in a position to consider the physical nature of these phenomena.

324. Rectilinear Propagation of Light—Shadows.—The fact that under ordinary circumstances light is propagated in straight lines is taken for granted by every one in common life, for we always assume that a body exists in the direction of the rays of light which enter our eye. The simplest proof of the rectilinear propagation is afforded by the formation of shadows, for it is found that the edge of the shadow of a body formed by a point source of light, such as a pin-hole in a dark

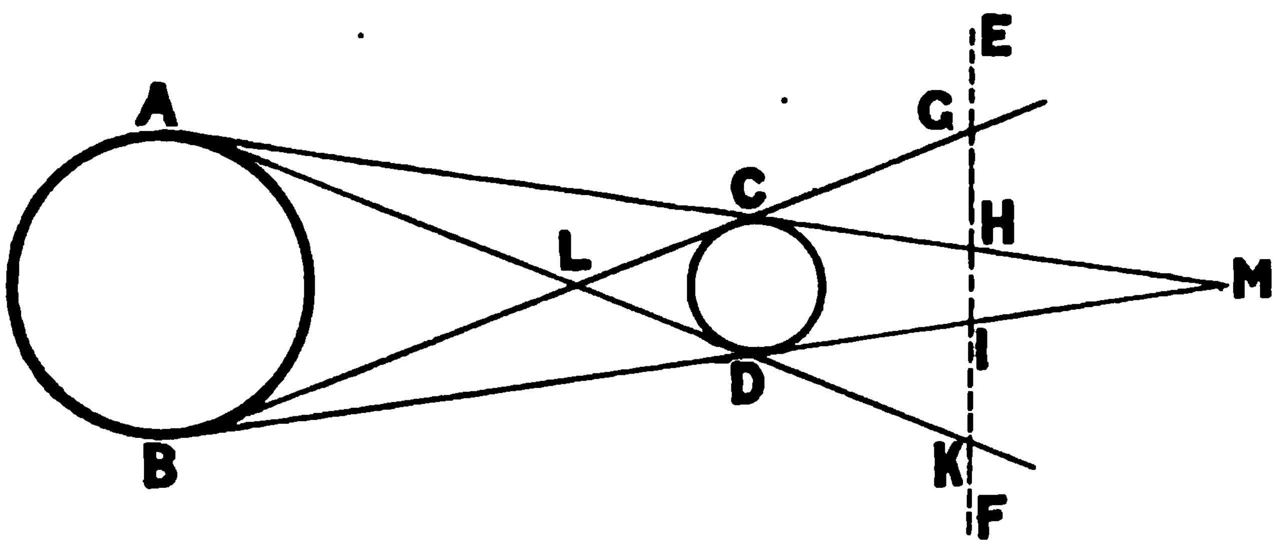


FIG. 280.

shutter, the edge of the object and the source of light are all in a straight line.

In order to find the form of the shadow cast by a point source, we have only to draw a number of straight lines from the source, so that they all touch the edge of the object. Where these lines meet the screen will be points on the edge of the shadow. If the source of light has an appreciable magnitude, however, we do not get a simple shadow of uniform blackness with a sharp outline. Let AB (Fig. 280) be a luminous object, say the sun, and CD the body that casts the shadow, say the moon. Then if we consider a point, A, of the luminous body, the shadow cast by this point on a screen at EF would be at HK. In the same way the shadow cast by the point B would be GI. All intermediate points would cast shadows situated between G and K. It will thus be seen that HI will be the only part of the screen which is completely in shadow, *i.e.* screened from the whole of the luminous object. This part of the shadow is therefore called the *umbra*. The rest of the shadow is not completely dark, but gets darker and darker from the outside to the

edge of the umbra. This part of the shadow is called the *penumbra*. In the case of the moon and earth, it is only when the earth enters within the cone CMD that a total eclipse takes place; when it enters within the penumbra the eclipse is only partial, since from any point within the penumbra straight lines can be drawn touching the object, which will intersect the source of light, and so part of the source will be visible from any such point.

325. The Pin-hole Camera.—The working of the pin-hole camera depends on the rectilinear propagation of light. If a small hole is made in an opaque screen, and a luminous object is placed on one side, and a white screen on the other, an inverted image of the luminous object will be formed on the screen. Each luminous point of the object A and B (Fig. 281) will form a small round patch of light on the screen; and if the hole is so small that these patches of light do not very much overlap, they will build up an image of the object, which, as is shown in the figure, is inverted. It is important to note that the image will be formed,

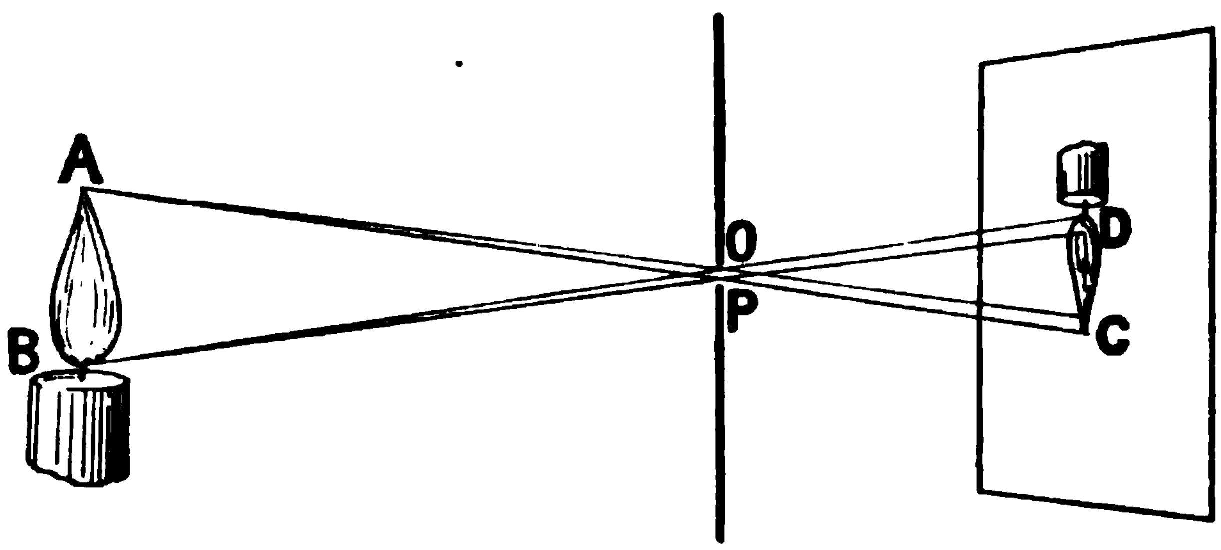


FIG. 281.

whatever may be the relative distance of the object and screen from the pin-hole, so that in this particular we have an important difference between the image formed in this way and that produced by a lens or mirror (§§ 337, 338). If a second pin-hole were made near the first, say at P, a second image would be produced, which would partly overlap the first image. In the same way, if a number of holes were made surrounding O, instead of a definite image we should simply have a blur produced by the partial superposition of all the images. This explains why it is that it is only when the pin-hole is small that any sharp image is obtained, for a large hole is the equivalent of a number of pin-holes close together.

326. Assumptions as to the Nature of Light.—We shall in a subsequent section describe experiments to prove that light travels with a finite velocity, and others which show that it is of the nature of a wave-motion. Since, however, for the full grasping of these experiments a knowledge of the laws of reflection and refraction and of the elementary properties of mirrors and lenses is required, the description of them is for the present postponed. We shall, nevertheless, in the following sections

assume that light consists of a wave-motion, and that the velocity with which the light waves move is different in different media, and on these assumptions we shall construct explanations of the simple phenomena of reflection and refraction. Thus when considering the properties of mirrors and lenses we shall not only use the method of the older geometrical optics, namely, the method of rays, but in addition we shall sometimes take as our starting-point the wave-front (§ 272) at any given instant, and then by Huyghens's construction (§ 273) we shall trace out the form of the wave-front at subsequent times. These two methods of viewing the phenomena are essentially the same, for the rays are everywhere at right angles to the corresponding wave-fronts; but it is nevertheless of use when employing the method of rays to have in our mind's eye the corresponding wave-fronts.

327. Curvature of a Surface.—If we have a disturbance produced at a point within an isotropic medium, the wave-fronts will be spheres with the point as centre. If, however, the medium is not isotropic, the form of the wave-fronts will in general be different. In the following pages we shall almost exclusively deal with spherical or plane wave-fronts. If we have a surface which is not a sphere, and at any point on this surface draw a sphere touching this surface, then for parts of the surface in the immediate neighbourhood of this point we may suppose the surface replaced by that of the sphere. In the same way we can draw a circle to touch any plane curve, and for points near the point of contact the circle will coincide with the curve.

Let AB (Fig. 282) be a portion of a curve, and at the points A and B draw two tangents to the curve. Now at the point A the direction of the curve is that of the tangent AT_1, while at B the direction of the curve is BT_2. Hence, when we pass from A to B the direction of the curve changes by an angle ADT_2 or θ. Let the length of the curve between A and B be S, then the rate of change in direction with distance measured along the curve is θ/S, and this is called the *curvature* of the curve between A and B. If the curve is a circle with its centre at C, the angle ACB is equal to the angles T_1DT_2, for the radii are at right angles to the tangents. Let r be the radius of the circle, then the length of the arc AB is $r\theta$ (§ 14). Hence the curvature of the circle is—

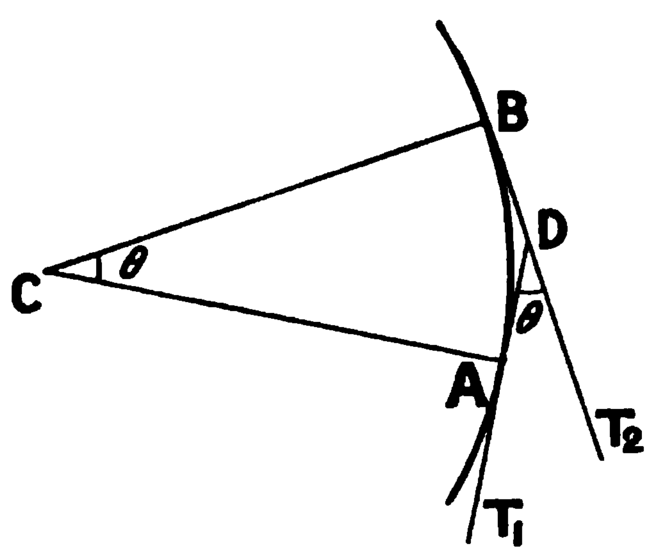


FIG. 282.

$$\theta/r\theta \text{ or } 1/r.$$

Thus the curvature of a circle is numerically equal to the reciprocal of the radius.

In the case of any other curve, if the tangent circle is drawn at any point the curvature of this circle is the reciprocal of the radius, and as the circle coincides with the curve at the given point, this also measures the curvature of the curve in the immediate neighbourhood of this point.

When in the place of plane curves we are dealing with surfaces, the same method is employed to measure the curvature, namely, the curvature at any point is equal to the reciprocal of the radius of the sphere which touches the surface at that point.

The radius of the tangent circle or sphere is called the *radius of curvature* of the curve or surface respectively, while the centre of the circle or sphere is called the *centre of curvature.*

In the case of a wave-front we have to distinguish two cases, namely, according as the direction in which the wave is moving is towards the concave or convex surface of the wave. We shall take the curvature of a wave to be positive when it is moving towards the centre of the tangent sphere. Thus the curvature of the spherical wave-surface produced by a disturbance at a point in an isotropic medium is negative.

328. Images.—If a wave has a positive curvature it is moving towards the centre of a circle, and if the medium between the wave-front and the centre of the circle is isotropic the wave will converge on this centre, so that at a certain instant the wave-front will be reduced to a point. Under these circumstances the wave is said to come to a real focus, or to produce a real image at the point.

If by reflection or refraction a wave-front of negative curvature is produced such that the centre of curvature does not coincide with the point where the wave was originated, the wave will travel as if it came from this centre, which is called a virtual focus, or virtual image.

Since the rays are always at right angles to the wave-fronts, a spherical wave-front of positive curvature corresponds to a pencil of rays which converge towards a point, this point being the centre of curvature of the wave-front. Thus a real image is produced when the rays of light which have started from a luminous point are, by reflection or refraction, caused to pass through a second point, this point being the real image.

In the same way, when the rays proceed as if they came from a point other than the actual source from which they do proceed, this point is called a virtual image.

In the case of a real image the waves and the rays actually pass through the image, while in the case of a virtual image the waves never actually pass through the image; they, however, proceed as if they had been produced at the image, and had then moved out in ever-widening spheres in an isotropic medium. In a virtual image also the rays never actually pass through the image; their direction, however, is such that if they were prolonged *backwards* they would pass through the virtual image.

329. Laws of Reflection.—The fact that bodies which are not themselves luminous are, when illuminated, visible in all directions,

shows that they must be capable of reflecting light in all directions, for it is by these reflected rays that we are able to see the body. Such rays, which are reflected from a body in all directions, and which often differ in many ways, such as colour, from the incident light, are said to have undergone diffused reflection.

When a beam of light is incident on a well-polished mirror, it undergoes diffused reflection to only a very small degree, the greater part of the light being reflected in a single direction. This light is said to have undergone regular reflection, and we now proceed to consider the laws that govern regular reflection.

The point where a ray of light strikes a mirror is called the point of incidence. If through the point of incidence a line, called the normal, is drawn at right angles to the reflecting surface, the angle the incident ray makes with this line is called the angle of incidence, while the angle the reflected ray makes with the normal is called the angle of reflection.

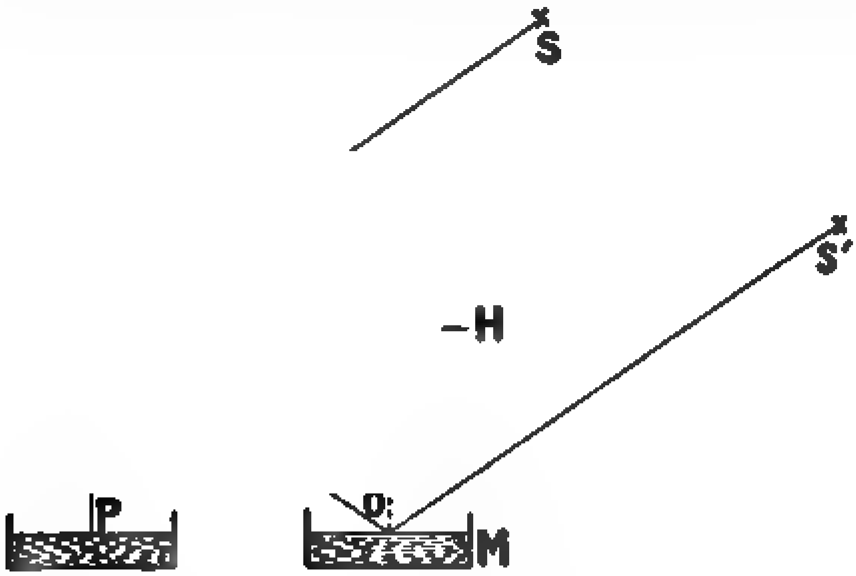


FIG. 283.

The phenomena of regular reflection may then be summed up in the following two laws :—

1. The incident ray, the normal to the reflecting surface at the point of incidence, and the reflected ray, are all in the same plane.

2. The angle of reflection is equal to the angle of incidence.

The laws of reflection are proved whenever an "artificial horizon" is used for determining the altitude of a star, *i.e.* the angle included by straight lines drawn in a vertical plane from the observer to the star and to the horizon. A telescope T (Fig. 283), movable in a vertical plane about a horizontal axle, is turned to observe a fixed star S, when seen directly, and the reading on a vertical divided circle is taken. The telescope is now turned down till the star is again seen in the telescope, but this time after reflection at the surface of some mercury in an open dish M. Since the surface of the mercury is horizontal, the normal at O is in

the vertical plane. As the telescope moves about a horizontal axle,[1] the plane in which it moves is vertical, and hence the two rays ST and OT are both in this vertical plane. The incident ray S′O is also in this plane, for, since the star is at such an enormous distance, the rays ST and S′O are parallel. Hence the first law is verified.

To prove the accuracy of the second law, a dish of mercury is placed immediately below the telescope, which is then turned till it is at right angles to the surface of the mercury. This adjustment is made by seeing when the image of the cross wires of the telescope seen reflected in the mercury surface coincides with the cross wires themselves. The circle attached to the telescope is then read, thus giving the reading when the telescope is pointing vertically downwards, *i.e.* normal to the surface of the mercury at P. Next, the circle is read when the image of the star, as seen directly and as seen after reflection in the mercury in M, coincides with the intersection of the cross wires.

From these readings the value of the angles P′TS and PTO can at once be deduced, and they are found to be equal. But since PP′ is a vertical, it is parallel to the normal ON, and therefore the angle PTO is equal to the angle TON, and the angle P′TS is equal to the angle S′ON, since ST is parallel to S′T. Hence the angle TON is equal to the angle S′ON, and the second law of reflection is verified.

330. Reflection at a Plane Surface.—We have in § 274 considered the reflection of a plane wave at a plane surface, and we have now to pass on to the cases when the incident wave is spherical, that is, is proceeding from a point disturbance, or, as it is generally called in optics, a point source or object.

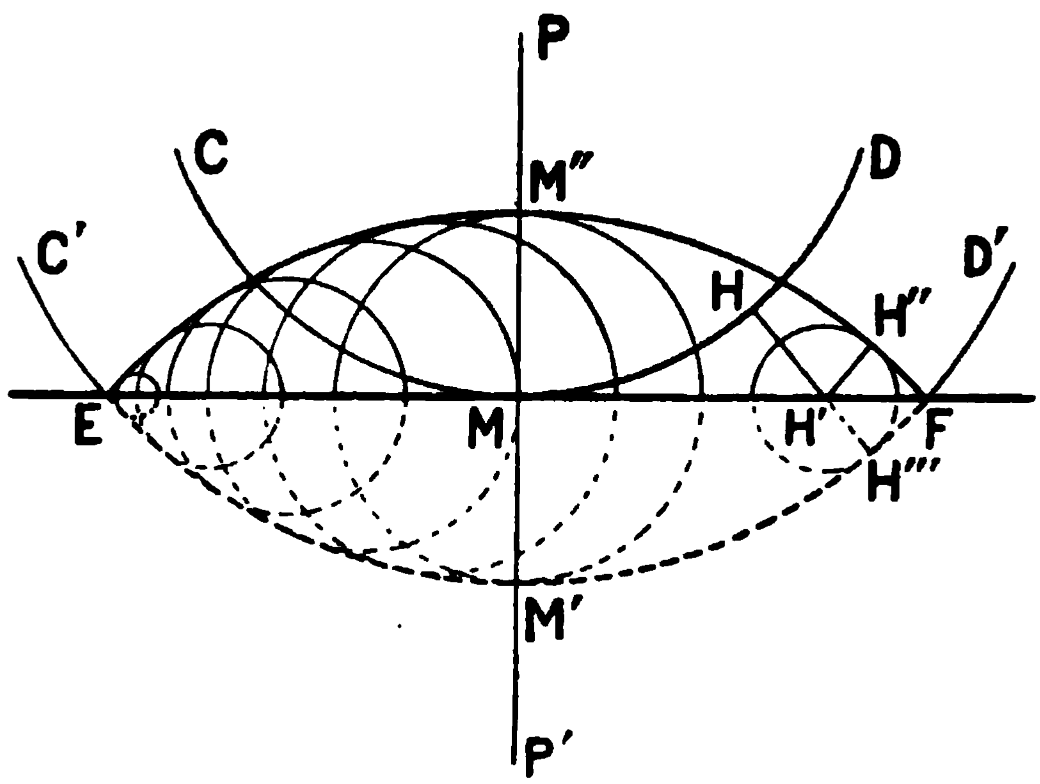


FIG. 284.

Let P (Fig. 284) be the luminous point, and EMF a section of the mirror. Let CMD be a wave-front which is just touching the mirror at M, and C′M′D′ be the position which the same wave-front would occupy at a time *t* later if no mirror were present. We may consider that as each portion of the incident wave-front reaches the mirror

[1] This condition is secured by means of a striding level which rests on the trunnions which carry the telescope.

it becomes the centre of a new disturbance which is propagated back into the medium, and then find the new wave-front by Huyghens's construction (§ 273). Thus, to find the position of the reflected wave at the time t, we describe from each point of EF a circle touching the arc EM′F, and the common tangent to all these circles will be the reflected wave-front. In the time t the reflected wave will travel from M through a distance equal to MM′, and it will reach the point M″, where MM″ is equal to MM′. In the same way the point H on the incident wave would, if there were no reflection, travel to H‴. It is, however, reflected at H′, and in the remainder of the time t is able to travel over a distance H′H″, which is equal to H′H‴. In the same way, it can be shown that the reflected disturbances are all circles touching the arc EM′F. Hence the reflected wave-surface EM″F is also an arc of a circle of the same radius as the circle C′M′D′. If P′ is the centre of the circle of which EM″F is an arc, P′M is equal to PM, for the radius P′M″ is equal to the radius PM′, and MM″ is equal to MM′. Thus P′ is the image of the point P produced by reflection in the mirror AB, and since the waves do not actually pass through P′, but only proceed after reflection as if they had originated at P′, the image is virtual.

Since EF touches the circle CMD, the line PM is perpendicular to AB. Thus the image P′ is situated on the line drawn from the luminous point at right angles to the mirror, and is as far behind the mirror as the object is in front.

The position of the image formed by reflection in a plane mirror can be deduced from the laws of reflection simply, without any assumption as to the nature of light.

Thus, as before, let AB (Fig. 285) represent the trace of a plane mirror which is perpendicular to the plane of the paper, and P be the luminous point from which rays of light are proceeding in all directions. Let PM_1 be a ray proceeding from P, which is reflected at M_1. The reflected ray will be along M_1Q_1, where the angles PM_1N, QM_1N are equal. In the same way, the rays PM_2 and PM_3 will be reflected along M_2Q_2 and M_3Q_3 respectively. Now, if the direction of these rays be produced backwards behind the mirror, it will be found that they all intersect at one point, P′. Thus, after reflection at the plane mirror AB, the rays of light proceed in the same directions that they would

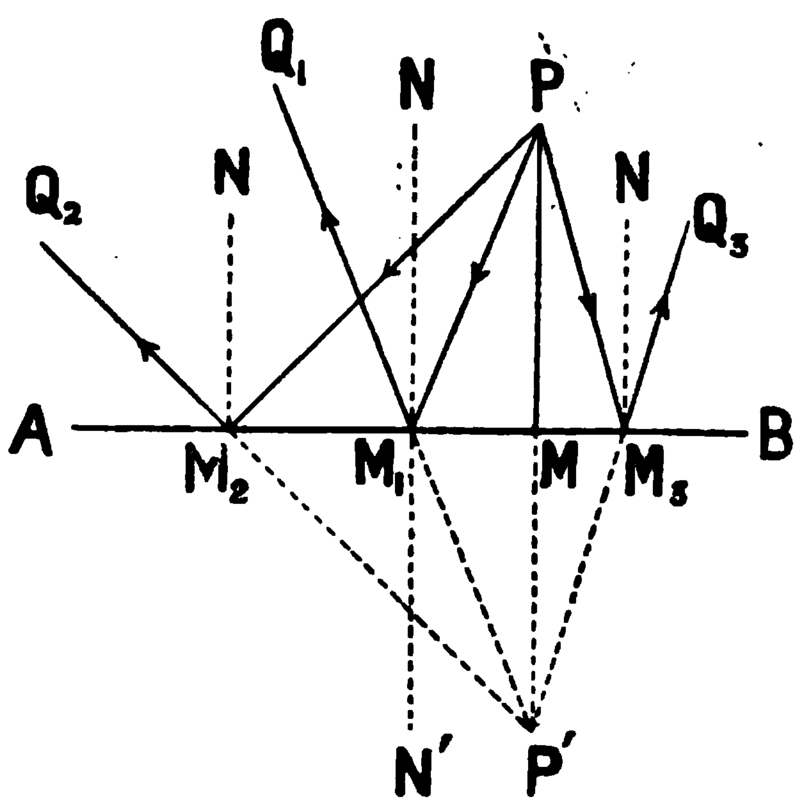


FIG. 285.

supposing the mirror were removed, and a luminous point were placed at P′. Thus the point P′ is the *image* of P formed by reflection in the mirror AB, and as the reflected rays do not actually pass through the image, but it is only their directions when produced backwards that pass through the image, the image is *virtual*.

In order to find the position of the image P′, we may proceed as follows: By the law of reflection the normal ray PM must be reflected back along the normal, and hence the line PMP′ must be normal to the mirror. By the same law the angle PM_1N is equal to the angle Q_1M_1N, which is itself equal to the angle $N'M_1P'$. Hence the angle PM_1M is equal to the angle $P'M_1M$. The angles PMM_1 and $P'MM_1$ are also equal, each being a right angle. The two triangles PM_1M and $P'M_1M$ have therefore one side common, and two angles of each equal; they are therefore equal in all respects, and the side PM is equal to the side P′M. The image P′ is therefore on the prolongation of the normal to the mirror drawn through the object P, and is as far behind the mirror as the object is in front.

In the case of a luminous object, as distinct from a point, which may either itself be a source of light, such as a candle flame, or simply appear luminous from the light which it reflects in a diffuse manner, each point produces its own image according to the above law. Thus the image of the arrow PQ (Fig. 286) in the plane mirror AB is at P′Q′. To an eye placed at E, the object as seen by reflection would appear to be at P′Q′, for, as shown, the pencils of rays which enter the eye diverge after reflection as if they came from the corresponding points of the image, and it is by the direction of the rays *as they enter the eye* that we judge of the position of a luminous body.

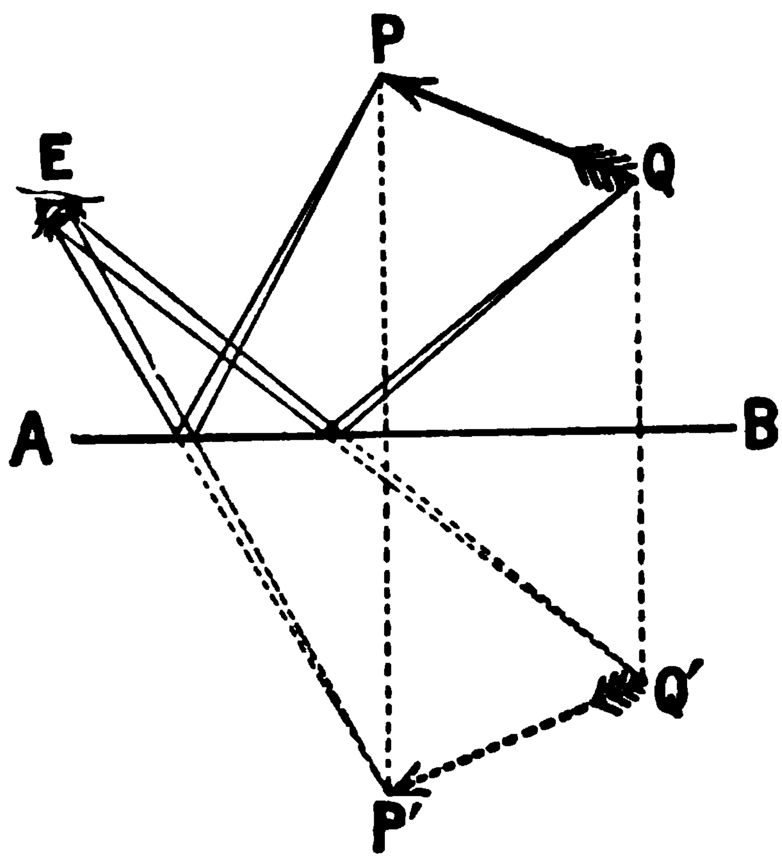


FIG. 286.

If from E the object PQ is viewed directly, the point of the arrow appears turned towards the left, while in the image the point (P′) appears turned towards the right; the image has therefore undergone perversion from right to left. This perversion by reflection is very clearly shown if some writing is blotted when wet on a clean piece of blotting-paper. In this way a perverted copy of the writing is obtained, and on holding this up before a plane mirror the image is perverted and the writing becomes at once legible.

331. Rotation of a Plane Mirror.—Suppose that a ray of light from the point P (Fig. 287) meets the plane mirror AB in the point M, and is reflected along MQ, and that the mirror is then rotated around an axis at right angles to the plane of incidence through an angle a into the position A′B′. The ray PM will now be reflected along MQ′, and we require to find the relation between the angle QMQ′ or β, through which the reflected ray has been rotated, and the angle a, through which the mirror has been rotated. Let MN and MN′ be the normals at the point M to the mirror in its two positions, then the angle NMN′ is equal to a. Also by the laws of reflection the angle PMN is equal to the angle NMQ, and the angle PMN′ is equal to the angle N′MQ′. Therefore the angle PMQ is equal to twice the angle PMN, and the angle PMQ′ is equal to twice the angle PMN′. Hence the angle β is equal to 2(PMN − PMN′). But the angle NMN′ or a is equal to PMN − PMN′. Hence $\beta = 2a$, or the reflected ray has been turned through twice the angle through which the mirror has been turned.

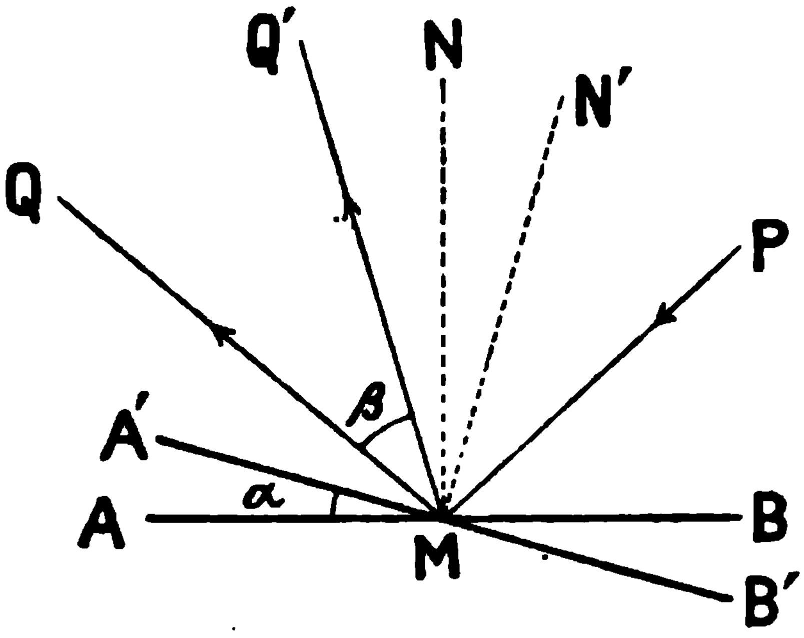


FIG. 287.

We may regard this problem in a slightly different manner. Let O (Fig. 288) be the point about which the mirror turns, then if a circle be described with O as centre and OP as radius, we shall show that whatever the position of the mirror, the image of P will lie on this circle. Let AB be a position of the mirror, then if P′ is the image of P, we have that PP′ is perpendicular to AB, and that PM = P′M. Hence in the two triangles OPM and OP′M, the two sides OM, MP of the one are equal to the two sides OM, MP′ of the other, and the included angles are equal, each being a right angle. Hence the triangles are equal in all respects, and OP is equal to OP′, hence P′ is on the circle described with O as centre and OP as radius. Let A′B′ be another position of the mirror, and P″ the image of P, then

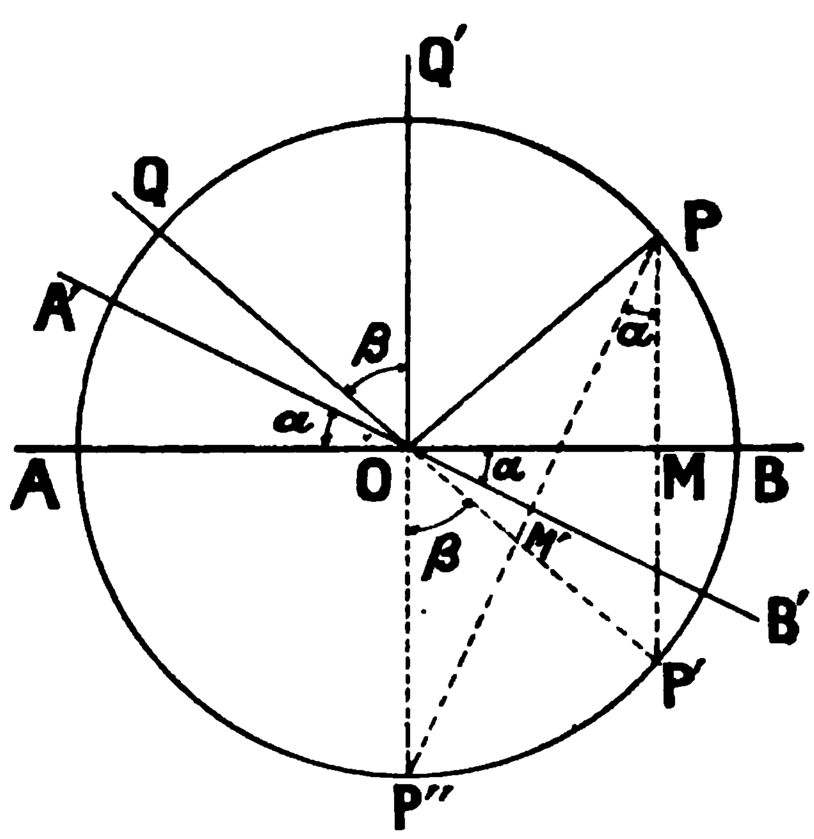


FIG. 288.

we can show just as before that P″O=OP, and hence P″ must be on the circle.

Since PM and PM′ are the normals to the mirror in the two positions, the angle M′PM or P″PP′ is equal to a, and by a well-known property of the circle, the angle P″OP′ subtended by the arc P″P′ at the centre is twice the angle P″PP′ subtended by the same arc at a point on the circumference. Hence the angle P″OP′ is $2a$, so that when the mirror turns through an angle a, the line joining the axis of rotation to the image turns through an angle $2a$. Now if PO is an incident ray, then OQ and OQ′ will be the reflected rays in the two positions of the mirror. Hence the angle QOQ′ included between the two reflected rays is $2a$.

332. Use of a Mirror and Scale to Measure an Angle.—Use is very frequently made of a mirror to measure the angle through which a body rotates, since we virtually get by this method a weightless unbendable pointer of whatever length we please.

Suppose ABB′ (Fig. 289) to be a scale divided, say, into millimetres, and that at a distance D from the scale is placed a mirror M, which can rotate about an axis perpendicular to the plane MAB′, *i.e.* the plane of the paper. Let MN be a line drawn from the centre of the mirror normal to the scale. If a ray of light is incident along PM, it will be reflected along MB when the mirror is in one position, and along MB′ when the mirror is rotated through an angle a, and we have seen that the angle BMB′ is equal to $2a$.

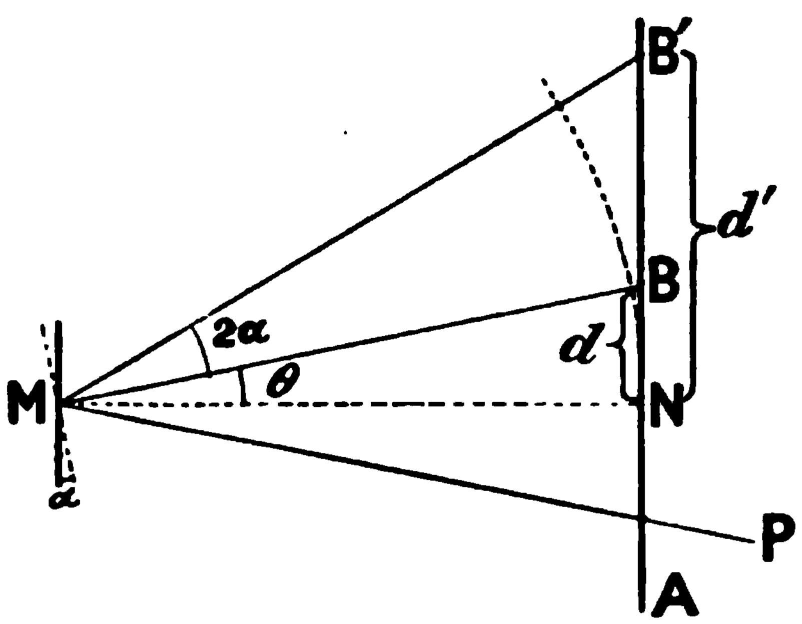


FIG. 289.

Calling the distance on the scale between the points B and B′, where the reflected rays cut the scale and the point N, d and d', and the angle BMN θ, we have :—

$$\frac{d}{D}=\tan\theta \text{ and } \frac{d'}{D}=\tan(2a+\theta).$$

Hence, knowing d, d', and D, we can from a table of tangents calculate the values of the angles θ and $2a+\theta$, and hence get a. Thus if the mirror M is attached to a body which can rotate, we can obtain the angle through which it rotates from the readings on the scale.

If, instead of being straight, the scale is curved so as to form part of a circle, having the mirror at its centre, then the difference of the readings for B and B′, divided by the radius D, will give the value of

the angle 2α in circular measure (§ 14) ; and hence, knowing the number of degrees in the unit angle in circular measure (radian), the angle α can be obtained in degrees.

If, when using the straight scale, the distance D between the mirror and the scale is very great compared to the lengths NB, NB′, both the angles θ and $\theta+2\alpha$ are very small. Hence, since in the case of very small angles the angle (in circular measure) is equal to the tangent (§ 14), we have

$$\frac{d}{D}=\theta \text{ and } \frac{d'}{D}=2\alpha+\theta.$$

Therefore
$$\alpha=\tfrac{1}{2}\,\frac{d'-d}{D},$$

or, if we express α in degrees,

$$\alpha^{\circ}=\frac{\pi}{360}\cdot\frac{d'-a'}{D}.$$

333. The Sextant.—The principle that the reflected ray is turned through twice the angle that the mirror is turned, is made use of in the sextant to measure the angle subtended at the observer by two distant objects, say the sun and the horizon. The sextant consists essentially of two mirrors, one of which, A (Fig. 290), is fixed, while the

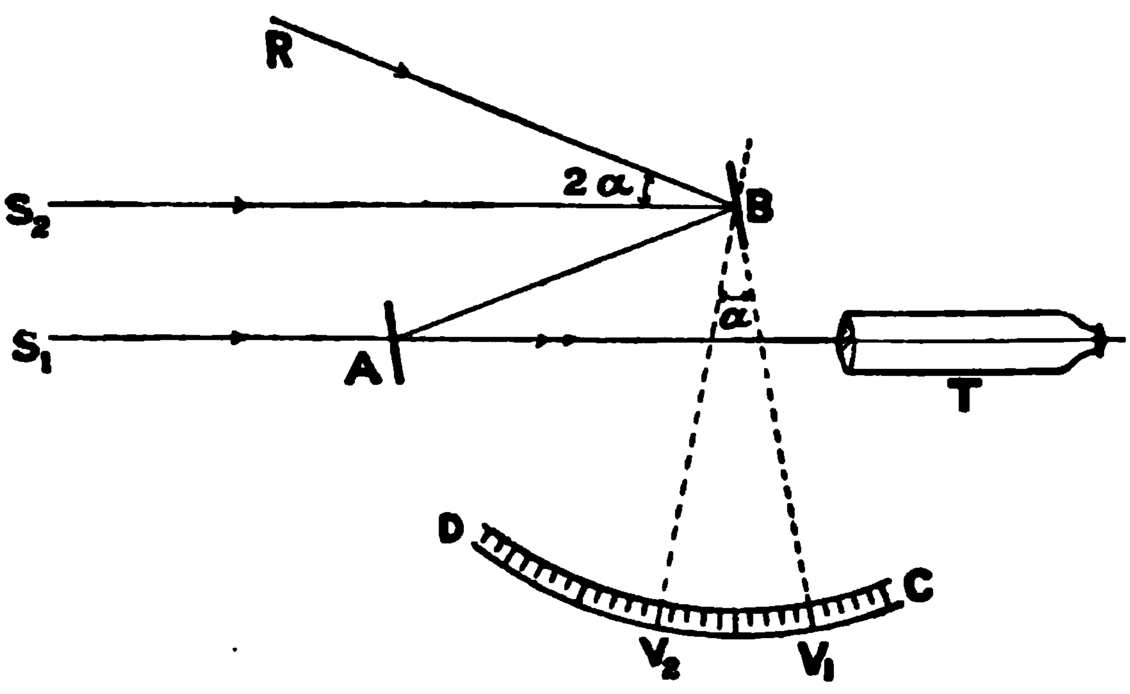


FIG. 290.

other, B, is movable about an axis at right angles to the paper, the angle through which it is turned being read by means of a vernier, V, and a graduated circular arc DC. The upper part of the mirror A has the silver removed, so that it is transparent, and a telescope T is so placed that half the light that enters the object-glass comes *through* the upper part of the mirror A, and the rest is light which has been reflected in the lower part. Suppose S_1A is a ray of light coming from a distant object, which traverses A and enters the telescope, and thus helps to produce an image of the object. Now the movable mirror B can be

so turned that a ray S_2B, coming from the same object, after reflection in the mirrors B and A, also enters the telescope. This ray will help to produce a second image of the object, which, by rotating the mirror B, can be brought alongside the image formed by the direct rays. Next, keeping the telescope turned so that the direct rays S_1 still form an image, turn the mirror B till the rays RB, proceeding from some other object, enter the telescope and form an image alongside the first. Since merely reversing the direction of the rays of light will not alter their paths, we now see that while in the first position of the mirror a ray incident along AB is reflected along BS_2, in the second position a ray incident along AB is reflected along BR. Hence, if a is the angle through which the mirror has been turned, the angle RBS_2 is equal to $2a$. Now the angle RBS_2 is the angle subtended at B by the two objects, and so, since a is obtained by reading the two positions of the vernier on the arc, this angle can at once be obtained. In order to save the necessity of doubling the reading of the scale, it is usual to number each half-degree as a whole degree, so that the reading gives directly the value of $2a$.

334. Reflection at Two Plane Mirrors.—If a luminous object P is placed in the angle included between two plane mirrors, AO and OB (Fig. 291), a *series* of images of P will be formed. In the first place we shall obtain two images, P_1 and P', formed by a single reflection in each of the mirrors. These images will, as proved in § 331, both lie on the circle drawn with O as centre and OP as radius. Suppose now that the mirror AO were removed and the object P replaced by its image, this will not affect the direction of the rays which, starting from P, after reflection at AO, strike the mirror OB. The point P_1 would then produce an image P'' in the mirror OB, which, since the line P_1P'' is perpendicular to OB, must be on the circle. This image is therefore formed by double reflection, first in AO and then in OB. In the same way, an image P_2 will be formed by reflection first in OB and then in AO.

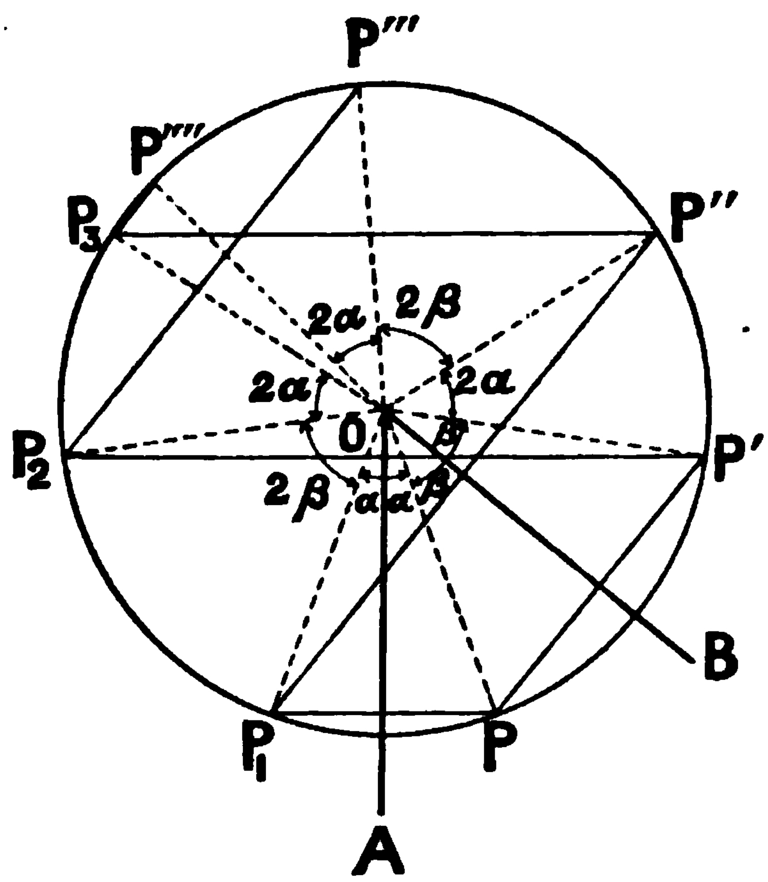


FIG. 291.

Again, the images P_2 and P'' will act as objects forming two new images, P''' and P_3, which are produced by reflection twice in one mirror and once in the other, and P_3 forms an image P'''' in OB produced by

reflection twice in OA and twice in OB. Neither P'''' nor P_3 can produce any more images, since these points are behind *both* mirrors, and you cannot get an image from a luminous point placed behind a mirror.

If we join OP, and call the angle AOP α, and the angle POB β, it can be shown from similar triangles that the angle BOP' is equal to β, and the angle AOP_1 is equal to α. Similarly, the angle BOP'' is equal to the angle BOP_1, that is, to $2\alpha+\beta$, so that the angle P'OP'' is equal to 2α. In the same way, the angle P_1OP_2 is equal to 2β. Also the angle P''OP''' is 2β, the angle P_2OP_3 is 2α, and the angle P'''OP'''' is 2α. It will thus be seen that the angles subtended by two consecutive images at the intersection of the mirrors are alternately equal to 2α and 2β.

An interesting case occurs when the angle between the mirrors is equal to $180°/n$, where n is some whole number. Since the angle between the mirrors is $\alpha+\beta$, we must have in this case $2n(\alpha+\beta)=360°$, that is, $2(\alpha+\beta)$ will divide exactly n times into the whole circumference. If, therefore, starting from P we mark off points on the circumference in both directions subtending alternately angles of 2α and 2β at the centre, we shall finally reach the same spot, so that the last image will be the image of the last but one of each of the two series, or, in other words, the two sets of images now have an image in common.

As an illustration, we may take the case when the mirrors are inclined at 60°, *i.e.* when $n=3$. If the object P (Fig. 292) is so placed that $\alpha=20°$ and $\beta=40°$, then we can obtain the images P'P''P''' by making the angle $POP'=2\beta=80°$, the angle $P'OP''=2\alpha=40°$, and the angle $P''OP'''=80°$; here we must stop, since P''' is behind both the mirrors. The whole angle POP''' is therefore $80+40+80=200°$. Proceeding in the same way, $POP_1 = 40°$, $P_1OP_2 = 80°$, and $P_2OP_3 = 40°$, so that the angle $POP_3=40+80+40=160°$. Hence the angles POP_3 and POP''' together amount to 360°, and therefore the points P_3 and P''' must coincide.

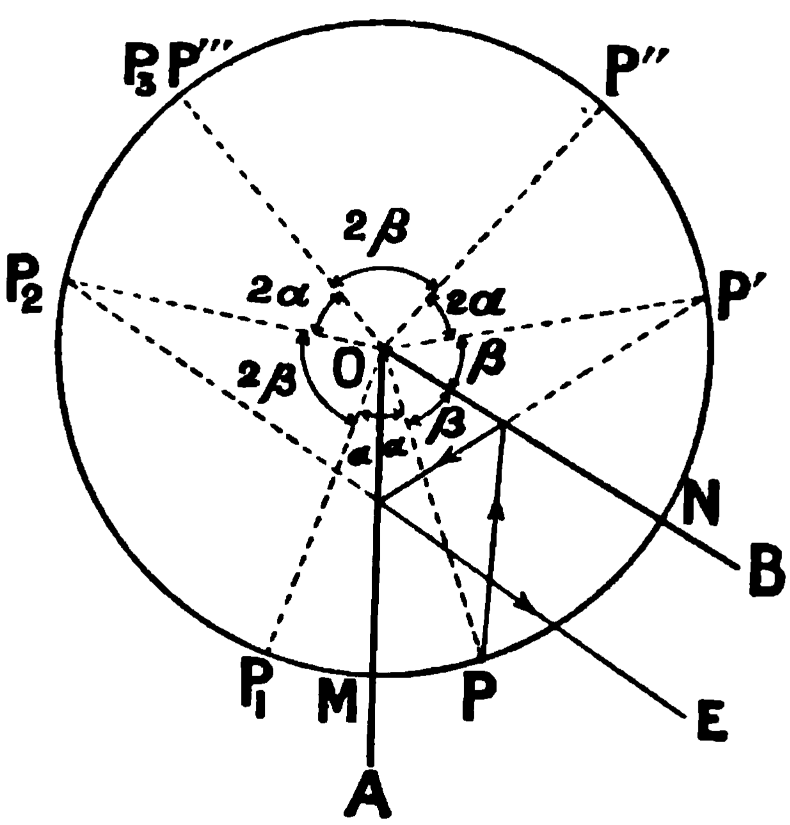


FIG. 292.

If we wish to trace out the actual path of a ray of light which, after leaving the object P, enters an eye at E, producing, say, the image P_2, we draw the line joining E and P_2, and join the point where this line cuts the mirror AO to the image P', since P_2 is the image in AO of P'. Finally, the point where this last line meets the mirror OB is joined to P. It will be easily seen that the line thus obtained is really

the path of the ray, for, starting from P, the ray after reflection in BO will travel as if it came from the image P′, and this ray, when it has been reflected in AO, will travel as if it came from P_2, *i.e.* along the line P_2E.

It is interesting to examine this same point from the view of the wave-fronts. Taking for simplicity the case where the mirrors AO, OB (Fig. 293) are at right angles, the circles drawn with the full line represent a series of wave-fronts which, starting from P, have not been reflected. When one of these wave-fronts meets the mirror AO it is reflected, and the circles with centre P_1, which are shown thus —‐‐—‐‐—, represent the reflected wave-fronts. In the same way, the circles shown thus — — — —

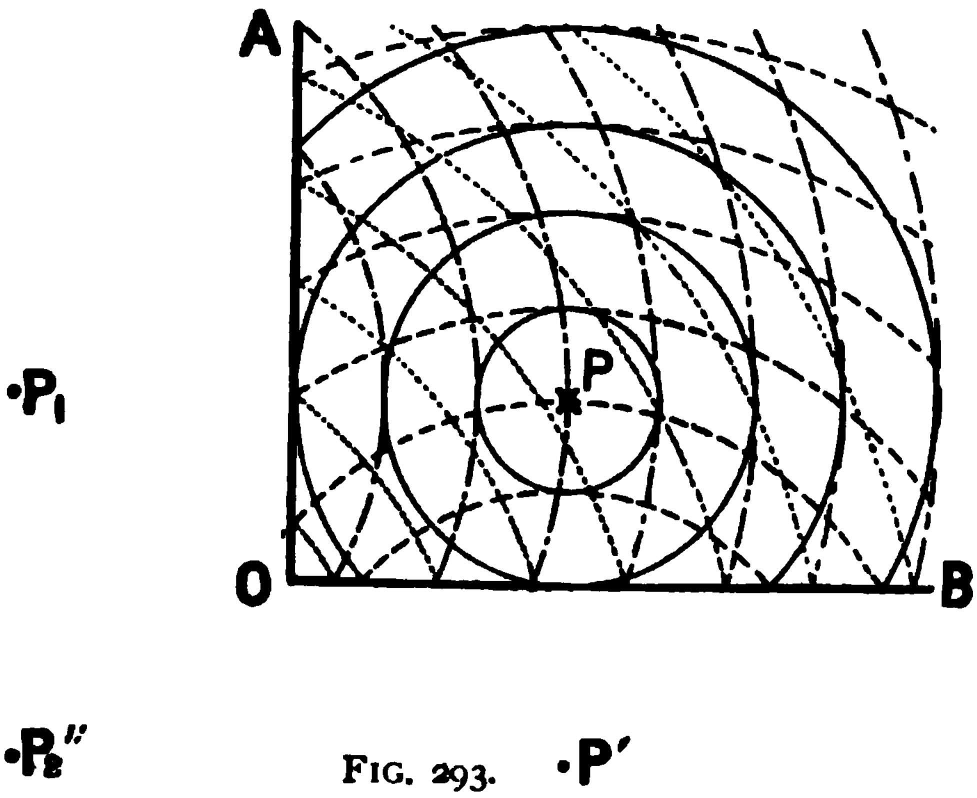


FIG. 293.

represent the wave-fronts as reflected in BO. When either of these reflected wave-fronts *impinges* on the other mirror it will be reflected, and in this way a series of wave-fronts with centre at P_2'' are obtained, and these are indicated thus .... It will be noticed how the circles with centre at P_2'' represent the reflected wave-fronts belonging to both sets of doubly reflected waves. This is of course because $90 = 180/2$, and so the second image of the two series coincide. An eye placed anywhere in the space included between the two mirrors will receive four sets of waves, of which the centres of curvature are P, P_1, P′, and P_2'' respectively, and will therefore see a luminous point apparently at each of these places.

The number of images formed if the angle between the mirrors is equal to $180°/n$ is $2n-1$, if the angle is less than $180°/n$ but greater than $\frac{180°}{n+1}$, the number of images is $2n$, since in the first case two of the images coincided.

As the angle between the mirrors is decreased the number of images is increased, and when the mirrors are parallel the angle between them is zero, and hence the expression $180°/n$ must be zero, which can only happen if n is infinite, that is, there must be an infinite number of images. Since the images all lie on a circle having its centre at the intersection of the mirrors and passing through the object, it is evident that, if the distance between the mirrors near the object is kept fixed, the radius of

the circle will increase as the angle between the mirrors is decreased. Finally, when the mirrors become parallel, their intersection is at an infinite distance, since parallel lines only meet at infinity, and therefore in this case the circle on which the images lie becomes one of infinite radius. Now if we consider a small part of a circle of very great radius, it is practically a straight line, which will be perpendicular to the radius of the circle at the part considered. Hence when the circle passing through P becomes of infinite radius, any finite portion of it near P will be a straight line, and this line will be perpendicular to the two mirrors, for the mirrors are radii of the circle. Hence we see that in the case of two parallel mirrors the images are infinite in number, and lie on a straight line drawn through P perpendicular to the mirrors. This could of course have been seen to be the case at once from first principles, but it is instructive to deduce it from the case of two mirrors inclined at a finite angle, since it is an easy illustration of the application of what is called the method of limits, a method of frequent use in physics.

We may also deduce the law according to which the images are spaced from the case of the inclined mirrors in the same way. We have, when the angle between the mirrors is finite, that the arc PM (Fig. 292) is equal to the arc P_1M, since they each subtend an angle a at O. Also the arc P_2P_3 and the arc $P'P''$, which each subtend an angle $2a$, are equal to twice the arc PM. In the same way, the arcs $P''P'''$ and P_1P_2 are each equal to twice the arc PN. Hence we might have said that the images are arranged round the circle so as to intercept arcs of the circumference which are alternately equal to 2PM and 2PN. When the circle becomes of infinite radius the arcs become portions of the same straight line, and hence the images are arranged along the line drawn through P normal to the mirrors, and at distances alternately equal to twice the distance between the object and the one mirror, and to twice the distance between the object and the other mirror. It will be a useful exercise for the student to deduce the positions of the images in the case of two parallel mirrors directly from first principles.

335. Multiple Images formed by a Thick Mirror.—If a ray of light proceeding from a luminous point P strikes a thick plate of glass ABCD (Fig. 294) which is silvered on the back CD at the point a, some of the light will be reflected from the surface of the glass along aa', forming an image of P

FIG. 294.

at P_1. The rest of the light will, however, penetrate into the glass travelling along aa'', and will be reflected at the silvered surface along $a''b$. At b part of this light will be reflected along bb'', and the rest will escape into the air and travel in the direction bb', forming an image at P_2. In the same way images will be formed at P_3, P_4, &c. Ordinarily the image P_2, formed by the light which has been once reflected at the silvered back of the mirror, will be much the brightest. If, however, the rays fall very obliquely, the amount of light reflected at a will be considerable, so that the image P_1 will be bright.

336. Measurement of the Angle of a Prism by Reflection.—Let AO and BO (Fig. 295) be two mirrors inclined at an angle, with their reflecting sides turned away from each other, and suppose we require to measure the angle AOB included between these reflecting surfaces. Let IO represent the direction of a parallel pencil of light, part of which is reflected in the mirror AO parallel to OR, and part is reflected in OB parallel to OR′. If ON is the normal at O to the mirror OB, then the angle of incidence ION is equal to the angle of reflection NOR′. If BO is produced to B′ and IO to C, then the angle IOB′ is equal to the angle COB. Also, since the angle ION is equal to the angle NOR′, the angle IOB′ must be equal to the angle BOR′, for the whole angles B′ON, BON are each right angles. Hence the angle COB is equal to the angle BOR′. In the same way, the angle COA is equal to the angle AOR, so that the angles AOC and COB are together equal to the angles AOC and COB, and therefore the angle ROR′, between the two reflected rays, is equal to twice the angle of the prism AOB. We shall see later on (§ 357) how the angle ROR′ may be measured.

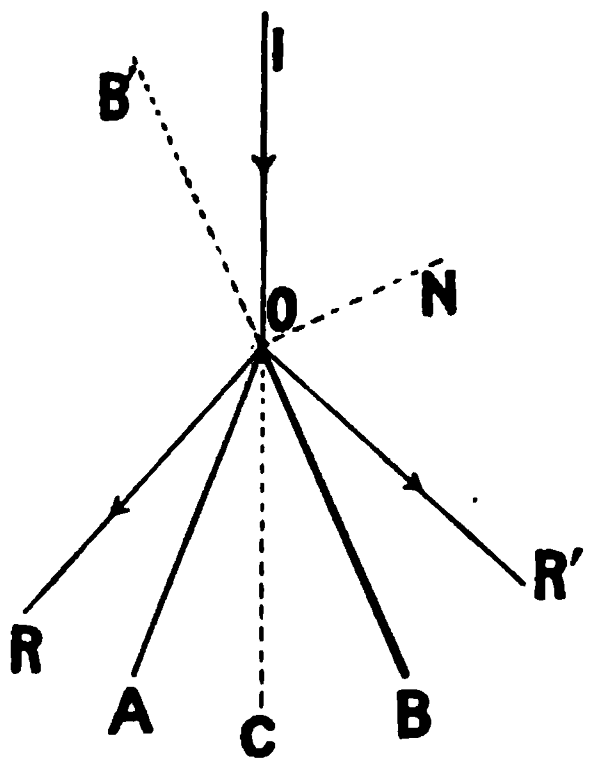


FIG. 295.

337. Reflection in Spherical Mirrors.—The smaller segment of a spherical reflecting surface cut off by a plane is called a spherical mirror. Spherical mirrors are of two kinds: (1) if the reflection occurs from the outside of the spherical surface the mirror is said to be *convex;* (2) if the reflection takes place from the inside of the spherical surface the mirror is said to be *concave.*

In order to investigate the reflection of light in spherical mirrors, we make use of the laws of reflection as given for a plane surface, for if we consider a very small element of surface on a spherical mirror, this small element will possess no appreciable curvature, so that we may treat it as a small plane mirror. The normal to the surface will at every point be

the radius of the sphere passing through the point, and hence the difference between the case of a spherical mirror and a plane mirror is, that while in a plane mirror the normals at all points are parallel, in a spherical mirror the normals are not parallel; they all, however, pass through a certain point, namely, the centre of the sphere of which the surface of the mirror forms a part.

A line drawn through the centre of the sphere of which a mirror is the part, perpendicular to the plane by which the mirror is cut off from the sphere, is called the *axis* of the mirror.

Let LOL′ (Fig. 296) represent a section of a spherical mirror, OC being the axis of the mirror and C the centre of the sphere from which the mirror is cut. The point C is called the *centre of curvature* of the mirror,

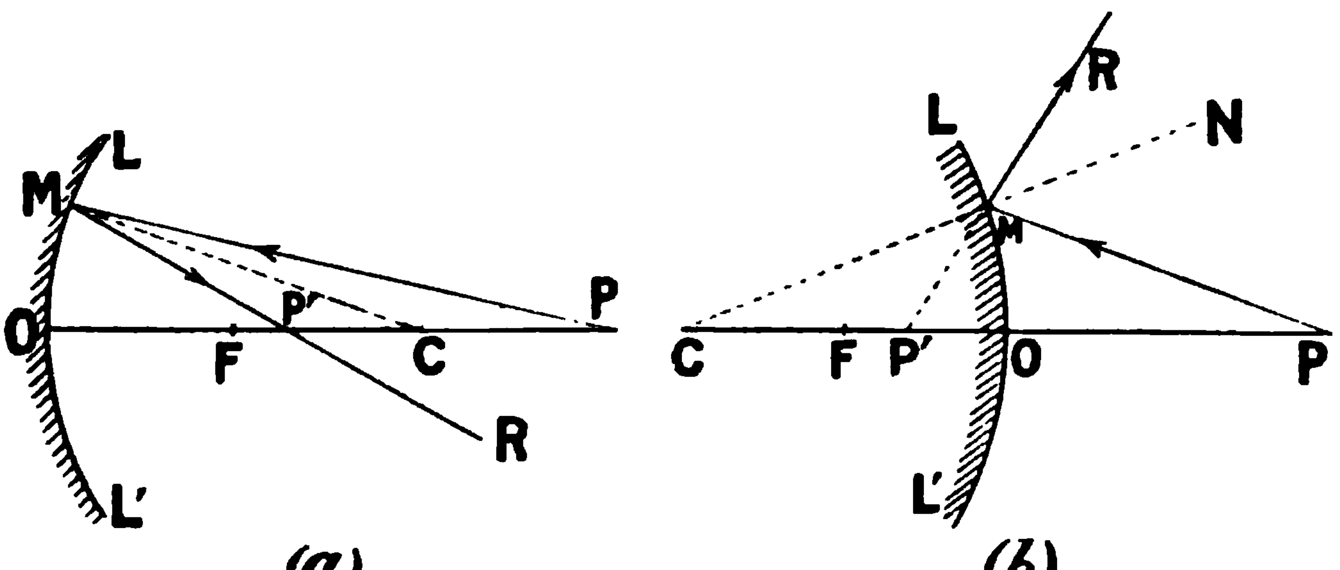


FIG. 296.

and the distance $\overline{OC}$, which is the radius of the sphere, is called the radius of curvature of the mirror.

Suppose that a luminous point P is placed on the axis of the mirror, then the ray of light from P incident along the axis will be reflected straight back, for the line PCO is normal to the mirror at O. If we join the point of incidence M, where another ray PM strikes the mirror, to the centre C, then the line CM in Fig. 296 (*a*), or CM produced in Fig. 296 (*b*), will be the normal to the mirror at M; and hence if we make the angle RMC equal to the angle PMC, or, in the case of the convex mirror, the angle RMN equal to the angle PMN, the line MR will be the path of the reflected ray. Let PR, or in the case of a convex mirror PR produced backwards, cut the axis at P′, then P′ will be the point of intersection of the two rays incident at O and M respectively after reflection.

Taking the case of the concave mirror first, we have. since in the triangle P′MP the line CM bisects the angle P′MP, the following relation (Euclid, vi. 3)—

$$\frac{\overline{CP}}{\overline{CP'}}=\frac{\overline{PM}}{\overline{P'M}}.$$

If the angle LCL′ subtended at the centre by the mirror, called the

aperture of the mirror, is small, then $\overline{PM}$ is very nearly equal to $\overline{PO}$, and $\overline{P'M}$ very nearly equal to $\overline{P'O}$, and under these circumstances

$$\frac{\overline{CP}}{\overline{CP'}}=\frac{\overline{PO}}{\overline{P'O}} \quad . \quad . \quad . \quad (1).$$

Thus, if the mirror is of small aperture, the position of P′ does not depend on the position of the point of incidence M, but only on the distance $\overline{PO}$ and the radius of curvature $\overline{OC}$ of the mirror. Hence all the reflected rays will pass through P′, and P′ will be the image of P produced by reflection in the mirror. We shall for the present confine our attention to mirrors of such small aperture that the assumptions made above hold good, so that P′ will be the image of the point P formed by reflection in the mirror, and since the reflected rays actually pass through P, the image is real.

We have now to make some convention as to the direction we shall call positive, and shall take all distances measured *from the mirror* in an *opposite* sense to that in which the incident light falls upon the mirror as *positive*, while all distances measured in the same sense as the incident light we shall take as *negative.*

Thus the distance $\overrightarrow{OC}$, Fig. 296 (*a*), being measured in the opposite sense to the incident light, which proceeds from P to the mirror in the sense $\overrightarrow{PO}$, is positive, while the distance $\overrightarrow{OC}$, Fig. 296 (*b*), is negative. It will also be convenient to use single letters to represent some of the distances which continually occur. We shall, therefore, in future indicate the radius of curvature $\overline{OC}$ of the mirror by r, the distance $\overline{OP}$ of the *object* from the mirror by u, and the distance $\overline{OP'}$ of the *image* from the mirror by v.

Now $\overline{CP}=\overline{OP}-\overline{OC}=u-r$; and $\overline{CP'}=\overline{OC}-\overline{OP'}=r-v$; hence the equation (1) reduces to

$$\frac{u-r}{r-v}=\frac{u}{v},$$

or

$$uv-vr=ur-uv.$$

$$\therefore \quad 2uv=vr+ur,$$

and dividing all through by uvr, we get

$$\frac{2}{r}=\frac{1}{u}+\frac{1}{r} \quad . \quad . \quad . \quad (2).$$

This equation gives us the general relation between the distances of object and image from a concave mirror of small aperture in terms of the radius of curvature of the mirror.

Returning to the case of a convex mirror, Fig. 296 (*b*), it will be noticed that the reflected rays do not actually pass through the image P′, but only their directions, so that the image is virtual.

As in the case of the concave mirror, we have

$$\frac{\overline{CP}}{\overline{CP'}}=\frac{\overline{PM}}{\overline{P'M}},$$

and when the mirror is of small aperture, this reduces to

$$\frac{\overline{CP}}{\overline{CP'}}=\frac{\overline{PO}}{\overline{P'O}},$$

Now $\overline{CP}=\overline{OC}+\overline{OP}=-r+u$, since $\overline{OC}$ is equal to $-r$, for the distance is measured in the negative direction, and $\overline{OP}$ is u; also $\overline{CP'}=\overline{OC}-\overline{OP'}$ $=-r+v$, both r and v being measured in the negative direction. Hence

$$\frac{u-r}{v-r}=\frac{u}{-v},$$

or

$$-uv+vr=uv-ur.$$

$$\therefore \quad 2uv=vr+ur.$$

$$\therefore \quad \frac{2}{r}=\frac{1}{u}+\frac{1}{r}.$$

Thus we have the same equation as in the case of the concave mirror. In making any numerical application of this formula it must, however, be carefully borne in mind that while for both classes of mirrors u is always positive, in the case of concave mirrors v and r are positive, while in the case of convex mirrors v and r are negative.

If we make the distance, u, of the object from the mirror larger and larger till the object is at an infinite distance, $1/u$ will become zero. Hence, under these circumstances,

$$v=\frac{r}{2},$$

and hence the image is formed at a point half-way between the mirror and the centre of curvature.

Since, when the object is at an infinite distance, all rays proceeding from it which strike the mirror may be considered as parallel, a pencil of parallel rays incident parallel to the axis are brought to a focus at a point on the axis at a distance equal to half the radius of curvature from the mirror. This point is called the *principal focus*, and the distance between it and the mirror is called the *focal length* of the mirror. In the case of a mirror the focal length (f) is equal to half the radius of curvature. Hence, in terms of the focal length, the formula for giving the position of the image becomes—

$$\frac{1}{f}=\frac{1}{u}+\frac{1}{v} \quad . \quad . \quad . \quad (3).$$

Since the principal focus is half-way between the mirror and the centre of curvature, the sign of f is the same as that of r, that is, f is positive for concave mirrors and negative for convex mirrors.

It is sometimes convenient to measure the distances of the object and image from the principal focus, instead of from the mirror. From Fig. 296 it will be at once seen that if the distance between the object and the principal focus F is called x, and the distance between the image and F is called y, these quantities being taken as positive if they are measured from F towards the object, we have for both kinds of mirrors, remembering that f is positive for concave mirrors and negative for convex mirrors—

$$u = x + f \text{ and } v = y + f.$$

Hence, substituting these values of u and v in equation (3), we get—

$$\frac{1}{f} = \frac{1}{x+f} + \frac{1}{y+f}$$

$$\therefore \quad (x+f)(y+f) = f(y+f) + f(x+f)$$

$$\therefore \quad xy = f^2 \quad . \quad . \quad . \quad (4).$$

This expression shows that the product of the distances of the object and image from the principal focus is equal to the square of the focal length. Since the square of any quantity must be positive, f^2 will always be positive, and hence the equation shows that the product xy must always be positive. Hence x and y must always be of the same sign, that is, the image and object must always lie on the same side of the principal focus.

Writing the expression (4) in the form

$$y = \frac{f^2}{x},$$

we can by giving x, the distance of the object, different values get the corresponding position of the image, as shown in the following tables :—

RELATIVE POSITIONS OF OBJECT AND IMAGE FOR A CONCAVE MIRROR.

Position of Object.	Value of x.	Corresponding Value of y.	Hence Position of Image.	Character of Image.		
At infinity .	∞	0	At principal focus ($p.f.$).	Real	...	...
Between ∞ and centre of curvature	+ & $>f$	+ & $<f$	Between $p.f.$ and centre of curvature	Real	Inverted	Diminished
At centre of curvature .	$+f$	$+f$	At centre of curvature .	Real	Inverted	Same size as object
Between centre of curvature and $p.f.$	+ & $<f$	+ & $>f$	Between centre of curvature and ∞	Real	Inverted	Magnified
At $p.f.$ . .	0	∞	At infinity .	...	...	...
Between $p.f.$ and mirror	− & $<f$	− & $>f$	Between mirror and $-\infty$	Virtual	Erect	Magnified
At surface of mirror . .	$-f$	$-f$	At surface of mirror . .	...	Erect	Same size as object

RELATIVE POSITIONS OF OBJECT AND IMAGE FOR A CONVEX MIRROR.

Position of Object.	Value of x.	Corresponding Value of y.	Hence Position of Image.	Character of Image.		
At infinity .	∞	0	At *p.f.* . .	Virtual	...	...
Between ∞ and the lens	$+$ & $>f$	$+$ & $<f$	Between *p.f.* and mirror	Virtual	Erect	Diminished
At surface of mirror . .	$+f$	$+f$	At surface of mirror . .	...	Erect	Same size as object

In the case of a convex mirror, x cannot be less than f.

338. Image of a Small Object on the Axis of a Mirror.—We have hitherto considered the image of a single luminous point, and now have to proceed to find the image of a small object placed on the axis of a mirror. Let $\overline{PQ}$ (Fig. 297) be such an object, then we may consider

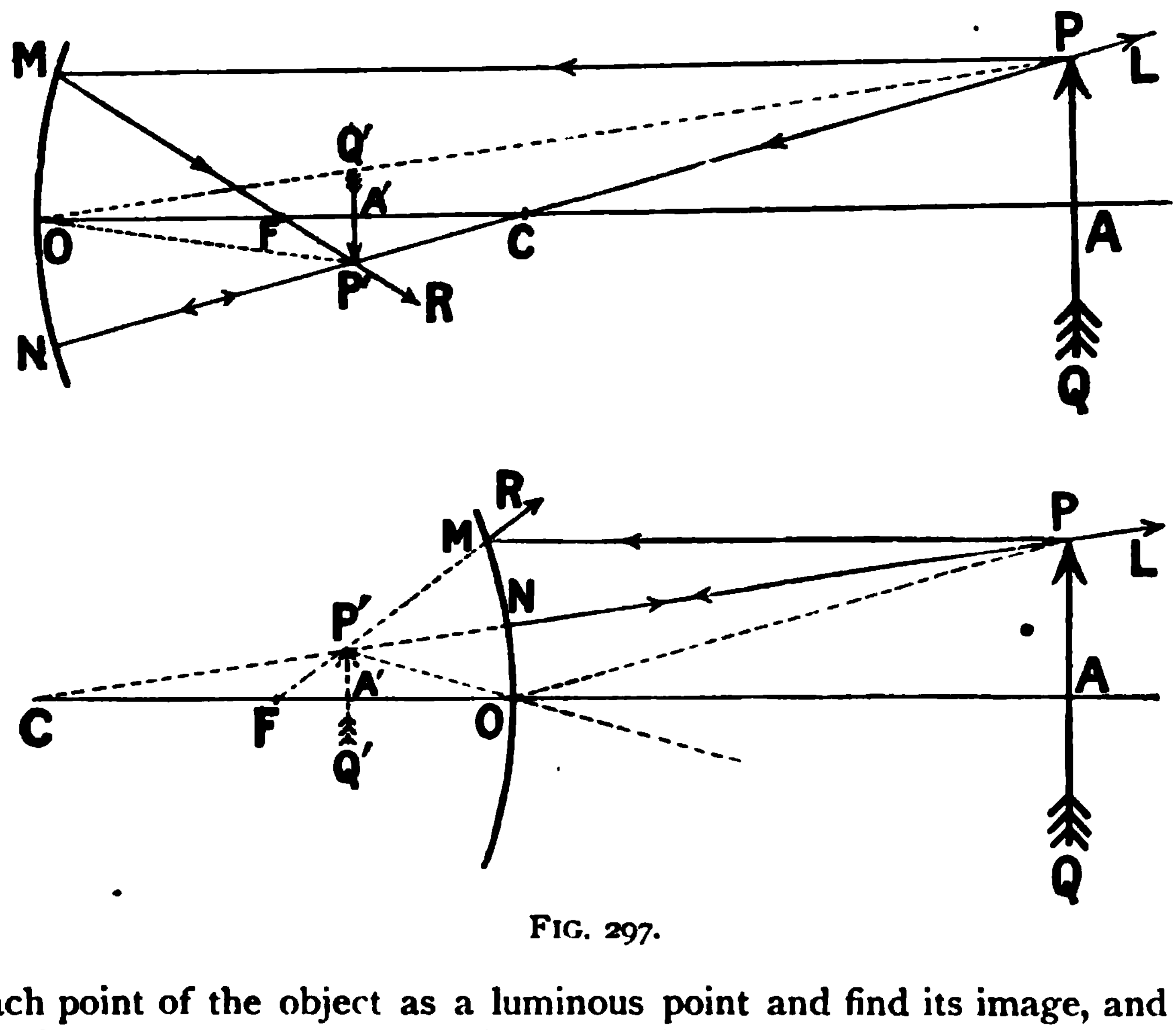


FIG. 297.

each point of the object as a luminous point and find its image, and all the images thus found will build up the image of the small object. The problem is most easily solved by a geometrical construction based on the results we have obtained in the last section.

Consider the point P of the object. A ray incident along the line PCN, passing through the centre of curvature of the mirror, will meet the

mirror normally at N and be reflected straight back along its path. A ray PM incident parallel to the axis of the mirror will, after reflection, either actually pass through the principal focus F (concave lens), or its direction when produced back will pass through the focus F (convex lens). These two rays, both proceeding from the point P, will therefore meet at P′, and this point will be the image of P. In the same way, the image of Q can be found by the intersection of one ray passing through the centre of curvature, which will be reflected back on itself, with another taken parallel to the axis, which, after reflection, will pass through the principal focus. We thus obtain the images of the extreme points of the object, and may fill in the intervening part free-hand, since the image and object will be similar. It will be seen that for the positions shown the images are in both cases smaller than the object, and that the image in the concave lens is inverted and real, while that in the convex lens is erect and virtual.

The relative sizes of the image and object can also be obtained from these figures, for a ray from P incident along PO will be reflected along OP′ to the image P′, and the angle of incidence POA must be equal to the angle of reflection AOP′. Hence the two triangles POA and P′OA′ are similar, and therefore

$$\frac{\overline{PA}}{\overline{P'A'}}=\frac{\overline{OA}}{\overline{OA'}}=\frac{u}{v}.$$

Also, since the triangles PCA and P′CA′ are similar,

$$\frac{\overline{PA}}{\overline{AC}}=\frac{\overline{P'A'}}{\overline{A'C}}.$$

Hence the ratio of the size of the object to that of the image is as the ratio of their distances from the mirror, or as the ratio of their distances from the centre of curvature.

The changes that take place in the relative size of object and image are given in the last two columns of the tables on pp. 462, 463.

339. Caustics formed by Reflection.—We have hitherto only considered reflection at spherical mirrors of such small aperture that all rays from a luminous point are reflected so that they pass through a single point. We now have to consider the directions of the reflected rays when the aperture of the mirror is large. A ray such as PM (Fig. 298), incident at a point M_1, near O, will be reflected so as to cut the axis at the point P′, which is the image of P given by the central part of the mirror. A ray such as PM_2, incident at a point M_2, at some distance from O, will not, however, be reflected through P′, but will intersect the axis nearer the mirror, at P″. In the same way the ray incident at M_3 will, after reflection, cut the axis at P‴. Hence the reflected rays will no longer all pass through a single point, this phenomenon being referred to as *spherical aberration.*

It is found that all the reflected rays are tangential to a certain curve EP′E′, which is called a *caustic curve*. Since near this curve the reflected rays will be more closely packed than at any other point, if a screen is placed so as to receive the reflected rays, the caustic curve will appear on the screen as a bright line. The caustic is also very clearly seen when a bright light shines on the inside of a cup nearly filled with milk, the surface of the milk acting as a screen.

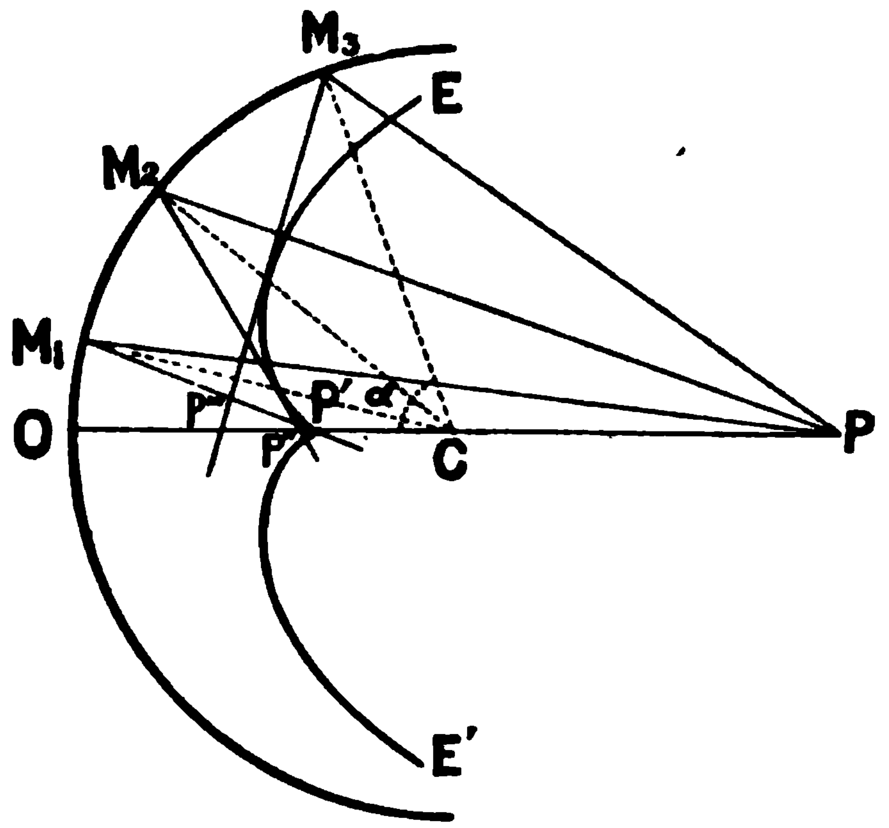


FIG. 298.

The caustic cuts the axis at the point P′, which is the image of P formed by the central parts of the mirror, and the curve forms a cusp at this point.

If we call the angle made by the normal, CM_3, at the point of incidence with the axis, α, then the distance of the point P‴, where a ray incident *parallel to the axis* when reflected cuts the axis from the point O, may be expressed by the equation

$$OP''' = r \times n,$$

where n is a fraction the value of which depends on the angle α.

The following table gives some values of n for different values of α :—

α.	n.
0°	·5
1°	·49994
5°	·49809
10°	·49229
20°	·45552

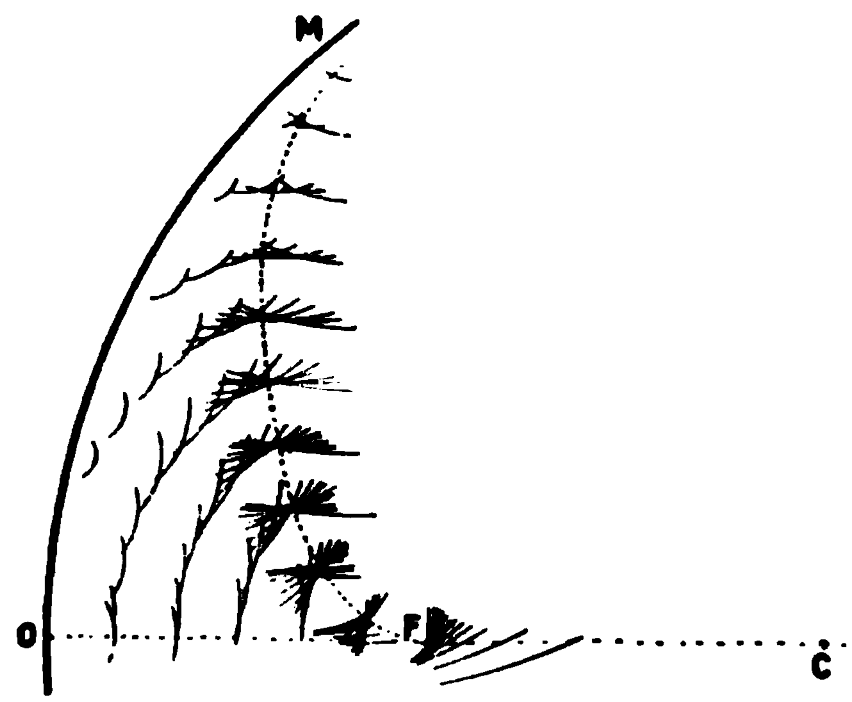


FIG. 299.

The manner in which the caustic is formed is very clearly shown in Fig. 299, where OM represents a section of a spherical mirror, of which OC is the axis and C the centre of curvature. The wave-fronts, after reflection in the mirror for a series of plane waves, incident parallel

to the axis, have been found by Huyghens' construction. It will be noticed how the circles representing the elementary waves are crowded together along the caustic, indicating that a violent disturbance is produced at all points on this curve.

340. Parabolic Mirrors.—Since in the case of spherical mirrors of any great aperture parallel rays are not all reflected through the principal focus, a luminous point placed at the principal focus will not produce, after reflection, a beam of parallel rays. As the brightness, and hence the distance to which a beam of light can be projected by a reflector depends on making the rays parallel, otherwise they become scattered over a larger and larger area as the distance from the mirror increases, it is of some importance to see if a mirror cannot be produced of such a form that parallel rays are all reflected through a single point, however great the aperture.

The surface formed by rotating a parabola about its axis fulfils this condition, for in the parabola the distance of any point on the curve OM

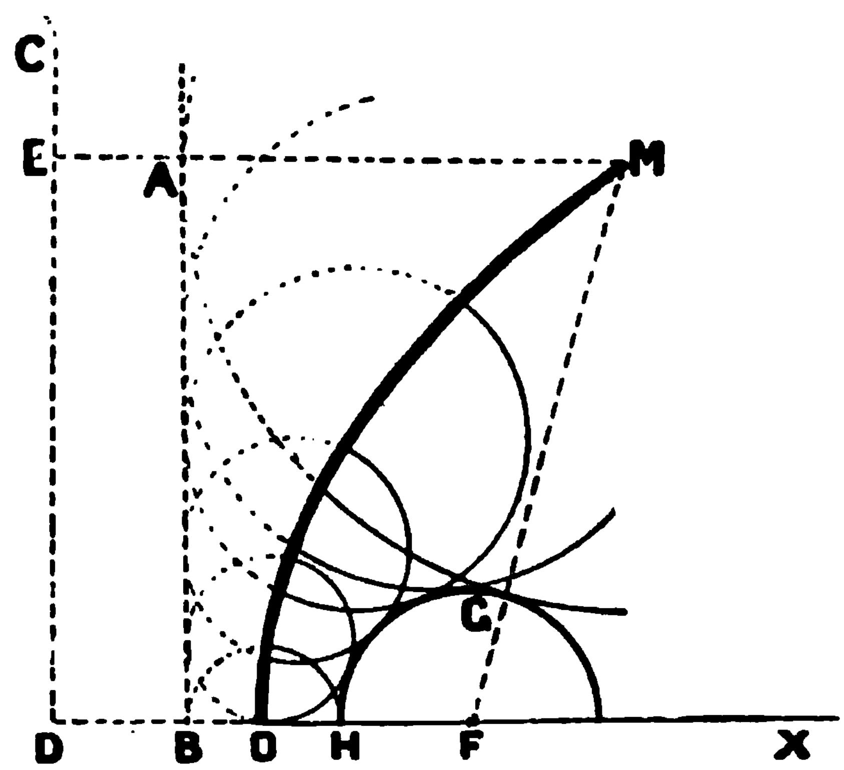


FIG. 300.

from the focus F (Fig. 300) is equal to its distance from the directrix CD. Hence if AB represents the position which a plane wave incident parallel to the axis would have occupied at a given instant, suppose the mirror MO were not present, we can obtain the position of the reflected wave-front by Huyghens's construction by drawing circles with their centres on OM, touching AB. These circles will all touch a circle HG described with the focus F as centre, for since by the property of the parabola $\overline{ME}=\overline{MF}$, and $\overline{MA}=\overline{MG}$, being both radii of the same circle, $\overline{FG}$ must be equal to $\overline{AE}$ or $\overline{BD}$. In the same way, it can be shown that the distance between F and any of the other circles is equal to $\overline{BD}$, so that a circle described with F as centre and $\overline{BD}$ as radius will touch all these

circles, and will represent the reflected wave-front. Since the section of the reflected wave-front is a circle, the wave will be brought to a focus at the centre of the circle, namely at F.

The difference between a spherical and a parabolic reflector, as far as the production of a plane wave, that is, a pencil of parallel rays, when a luminous point is placed at the focus, is shown in Fig. 301. The portion

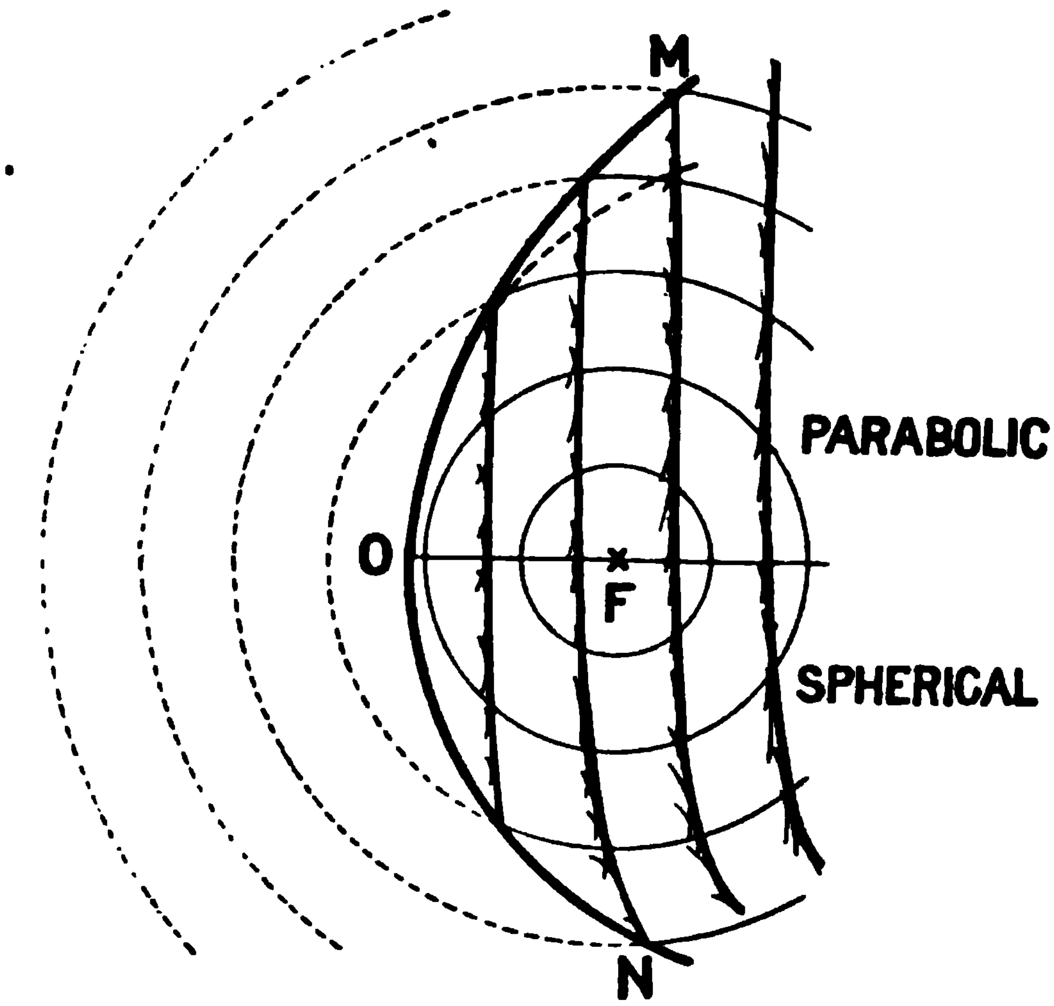


FIG. 301.

MO of the mirror is parabolic, while the portion ON is spherical. A luminous point is supposed to be placed at F, and the positions of a series of reflected wave-fronts have been drawn by Huyghens's construction. While the wave-fronts after reflection at the parabolic mirror are plane, this is only the case for those portions quite close to the axis, when reflection takes place from the spherical surface.

CHAPTER II

REFRACTION

341. Refraction—Snell's Law.—As long as a ray of light travels through a homogeneous (isotropic) medium, its path is a straight line; in general, however, when it passes from one medium to another, the direction of the path of the ray changes abruptly at the surface of separation of the two media. In addition to the portion of the light which penetrates into the second medium, a portion of the light will be reflected at the surface of separation, according to the laws we have just considered, and we shall, in the present section, generally neglect the consideration of this reflected ray, and concern ourselves exclusively with the portion which penetrates into the second medium, and which is called the *refracted ray*.

As before, we shall call the point where the incident ray meets the surface of separation between the two media the point of incidence, also the angle of incidence, the plane of incidence, and the normal will have the same signification as in the case of reflection. The angle between the refracted ray and the normal in the second medium will be called the *angle of refraction*.

We then have the following laws :—

(1) The refracted ray lies in the plane of incidence, and on the opposite side of the normal to the incident ray.

(2) The sine of the angle of refraction bears a constant ratio to the sine of the angle of incidence for all angles of incidence, the value of the ratio depending on the nature of the two media at the surface of separation between which the refraction takes place, and also on the nature of the incident light (Snell's law).

It will thus be seen that in the case of refraction the conditions are much more complicated than in that of reflection, for while in the latter the *direction* of the reflected ray was independent of the nature of the reflecting surface, of the medium in which the light was travelling, and of the nature (colour) of the light, in the case of refraction the direction of the refracted ray depends on all these conditions. We shall for the present postpone the consideration of the influence of the nature of the light, assuming that the light with which we are about to deal is the yellow light given out by a Bunsen flame when a bead of common salt (NaCl) is placed in the flame.

Suppose AB (Fig. 302) to be the surface of separation between two media, say air above and glass below, and that a ray of light travelling in the direction IO is incident at O. Let NON′ be the normal to the surface of separation at O, then in the case considered, in which the medium above AB is less dense than that below, the angle of refraction RON′, or β, will be less than the angle of incidence ION, or α. If the medium above AB had been denser than that below, then α would have been less than β.

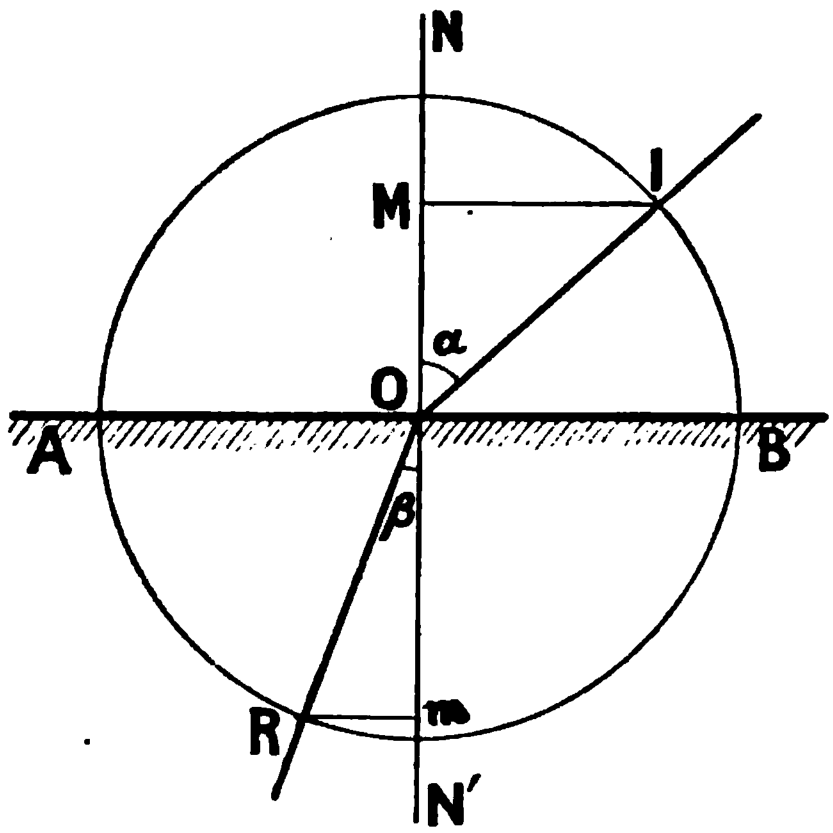


FIG. 302.

If, with O as centre, we describe a circle of any radius, cutting the incident ray at I and the refracted ray at R, and from I and R draw perpendiculars to the normal, then

$$\sin\alpha=\frac{\overline{IM}}{\overline{IO}}\text{; and }\sin\beta=\frac{\overline{Rm}}{\overline{RO}},$$

or

$$\frac{\sin\alpha}{\sin\beta}=\frac{\overline{IM}}{\overline{Rm}}.$$

According to Snell's law the ratio $\sin\alpha/\sin\beta$ is constant for all angles of incidence, and the value of this ratio for any pair of media is called the *refractive index* for these media, and is generally indicated by the Greek letter μ.

When the incident ray is perpendicular to the surface of separation α is zero, and hence $\sin\alpha=0$, so that $\sin\beta=0$ and $\beta=0$. Thus in this case the ray does not suffer refraction.

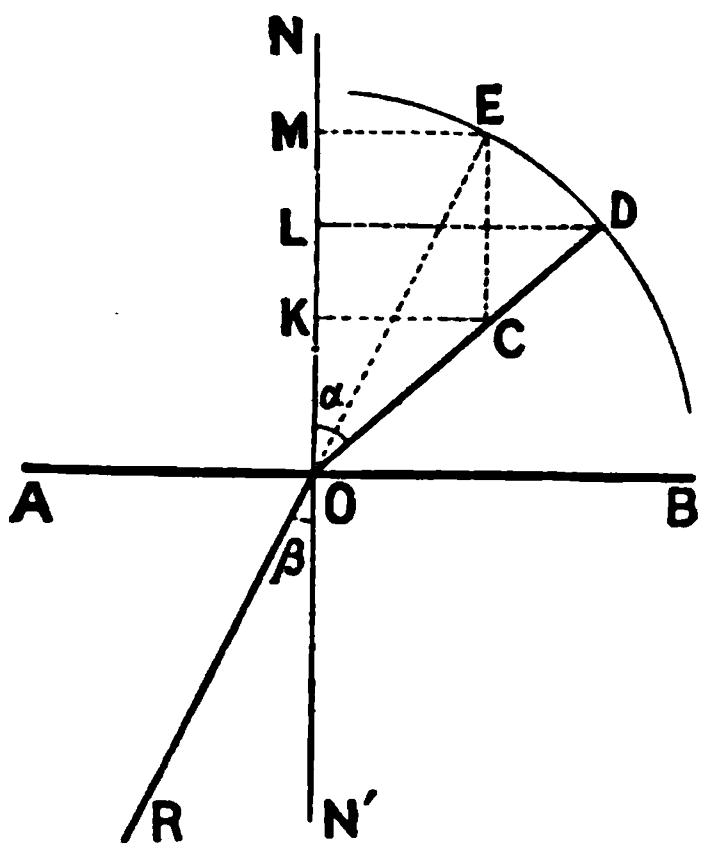


FIG. 303.

If we are given the refractive index between two media and the angle of incidence, it is easy, by a geometrical construction, to find the direction of the refracted ray. If AB (Fig. 303) is the surface of separation between the media, the denser being below, and DO is the direction of the incident ray, measure off from O along $\overline{OD}$ a distance OC equal to unity, or

with O as centre describe a circle of radius unity cutting DO at C, and a distance $\overline{OD}$, which, expressed in the same units, is equal to the refractive index μ. With centre O and radius $\overline{OD}$ describe an arc of a circle. From C draw CE parallel to the normal ON, cutting the circle in E, join EO and produce to R, then OR is the refracted ray.

To prove this, draw EM, DL, and CK perpendiculars to ON.

Now $$\frac{\overline{DL}}{\overline{OD}}=\sin\alpha,$$

and $$\frac{\overline{ME}}{\overline{OE}}=\sin\beta.$$

Hence $$\frac{\sin\alpha}{\sin\beta}=\frac{\overline{DL}}{\overline{OD}}\cdot\frac{\overline{OE}}{\overline{ME}}$$
$$=\frac{\overline{DL}}{\overline{KC}},$$

for $\overline{OD}$ and $\overline{OE}$ are equal, being radii of the circle, and $\overline{ME}$ is equal to KC. Now the triangles DOL and COK are similar. Hence

$$\frac{\overline{DL}}{\overline{KC}}=\frac{\overline{OD}}{\overline{OC}}.$$

But by construction $$\frac{\overline{OD}}{\overline{OC}}=\mu.$$

Hence $$\frac{\sin\alpha}{\sin\beta}=\mu,$$

and β is the angle of refraction, so that $\overline{OR}$ is the direction of the refracted ray.

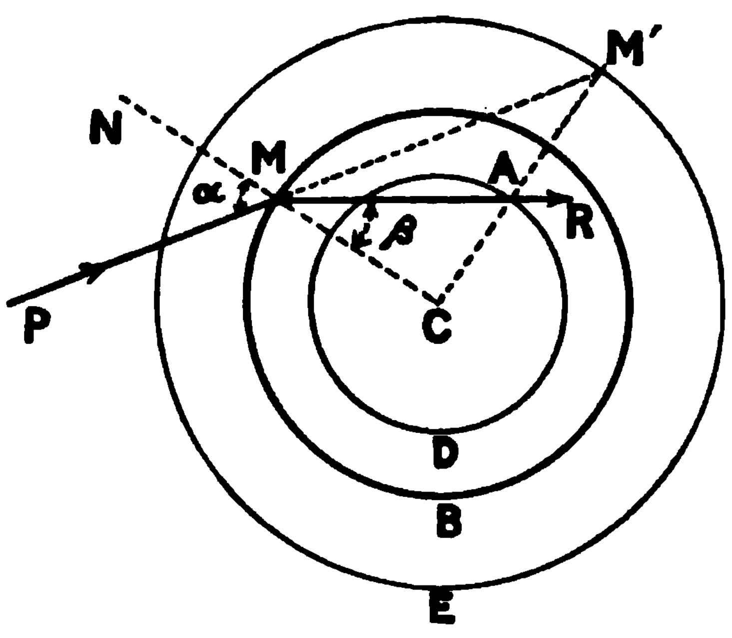


FIG. 304.

If, instead of being a plane, the surface separating the two media is a sphere of radius r, the path of the refracted ray may be found by the following geometrical construction. Let the circle MB (Fig. 304) with centre C be a meridian section of the sphere of the medium, the refractive index between the surrounding medium and the sphere being μ. With C as centre, describe two circles, having radii μr and r/μ respectively, and produce the incident ray PM till it cuts the outer circle at M′. Join M′C, cutting the inner circle at A. Then the line MAR will be the refracted ray.

By construction we have

$$\frac{\overline{MC}}{\overline{CA}}=\frac{r}{r/\mu}=\mu,$$

and

$$\frac{\overline{M'C}}{\overline{MC}}=\frac{r\mu}{r}=\mu.$$

Also, since in the triangles MCA and M'CM the angle MCA is common, and, as above shown, the sides about these angles are proportional, it follows that the triangles are similar. Hence the angle MAC is equal to the angle M'MC, or α.

In the triangle MCA, since the ratio of two sides is the same as the ratio of the opposite angles, we have

$$\frac{\overline{CM}}{\overline{CA}}=\frac{\sin MAC}{\sin CMA},$$

or

$$\frac{\mu r}{r}=\frac{\sin MAC}{\sin \beta}=\frac{\sin \alpha}{\sin \beta}.$$

But if MA is the refracted ray,

$$\frac{\sin \alpha}{\sin \beta}=\mu,$$

and since the line MA has been shown to fulfil this condition, it must be the refracted ray.

342. Refraction through a Slab with Parallel Sides.—Suppose we have a slab of a denser medium enclosed by parallel sides, AB and CD (Fig. 305), with a less dense medium on either side. Then it is found experimentally that if a ray of light is passed through the plate, the direction O_1R of the ray after leaving the plate is parallel to the incident direction IO, the only effect of the interposition of the plate being to displace the ray to one side.

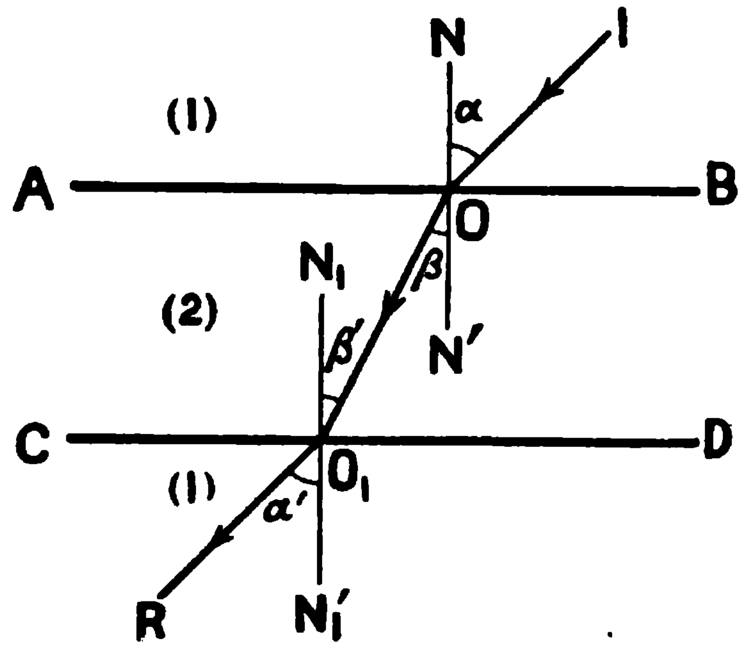


FIG. 305.

We will call the less dense medium 1, and the medium composing the slab 2, and indicate the refractive index from medium 1 to medium 2 by ${}_1\mu_2$, and that from 2 to 1 by ${}_2\mu_1$. We have

$${}_1\mu_2=\frac{\sin \alpha}{\sin \beta},$$

and

$${}_2\mu_1=\frac{\sin \beta'}{\sin \alpha'}.$$

Since the sides of the slab are parallel, and NON' and $N_1O_1N_1'$ are normals,

these lines are parallel, and therefore the angle β is equal to the angle β'; also, since the rays IO and O_1R are found by experiment to be parallel, the angle α is equal to the angle α'.

Hence $$\frac{\sin \alpha}{\sin \beta}=\frac{\sin \alpha'}{\sin \beta'},$$

$$\therefore \qquad {}_1\mu_2=\frac{1}{{}_2\mu_1},$$

or $${}_1\mu_2 \cdot {}_2\mu_1=1.$$

Thus we get that the refractive index from medium (1) into medium (2) is the reciprocal of the refractive index from medium (2) into medium (1).

By taking a number of slabs of media of different refrangibility, it can be shown, using a similar notation to that employed above, that

$${}_1\mu_2 \cdot {}_2\mu_3 \cdot {}_3\mu_4 \cdot {}_4\mu_5 \quad . \quad . \quad . \quad {}_n\mu_1=1.$$

This expression will be of use in solving problems on refractive index. Thus, given that the refractive index from air to glass is 1.5, and that from air to water is 1.33, find the refractive index from water to glass. In the first place,

$$\mu(\textit{air to glass}) \times \mu(\textit{glass to air})=1,$$

$$\therefore \qquad \mu\ (\textit{glass to air})=\frac{1}{1.5}=.67\,..$$

Also $$\mu(\textit{air to water}) \times \mu(\textit{water to glass}) \times \mu(\textit{glass to air})=1,$$

$$\therefore \qquad 1.33 \times \mu(\textit{water to glass}) \times .67=1.$$

Hence $$\mu(\textit{water to glass})=\frac{1}{1.33 \times .67}=1.13.$$

or the refractive index from water to glass is 1.13.

343. Image of a Point Formed by Refraction at a Plane Surface.—If AB (Fig. 306) is the surface separating two media, the refractive index from one to the other being μ and the denser below AB, and P is a luminous point in the denser medium, the ray PMN, which strikes the surface of separation normally, is unrefracted. All other rays, such as PM_1, which strike the surface AB obliquely will be refracted, the direction of the refracted ray M_1R_1 being obtained by the construction given in § 341. If we produce M_1R_1 backwards, it will intersect the normal ray PN at the point P′, and the refracted rays MN and M_1R_1 will proceed as if they came from P′. If we make the same construction for a ray such as PM_2, which strikes the surface AB a good deal further from M, the refracted ray M_2R_2, when produced backwards, will be found to intersect the normal ray at a point P″, nearer to M than P′. Hence the directions of the refracted rays do not all pass through a single point,

so that there is not a single geometrical image of P. The directions of all the refracted rays are however tangential to a caustic curve CP′.

If we restrict ourselves to rays which strike the surface near any given point M, a restriction similar to that made with regard to spherical mirrors, it will be found that the directions of all the refracted rays very nearly pass through a single point, so that in this case we get an image of the luminous point. If we draw the normal at M (Fig. 307), and call the angle N″MR α, and the angle PMN′ β, we have

$$\frac{\sin \alpha}{\sin \beta} = \mu$$

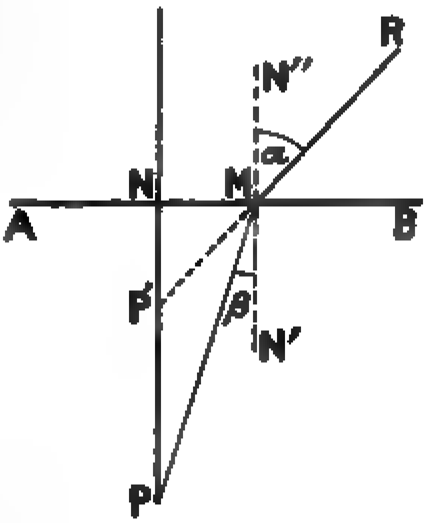


FIG. 307.

Now since the angles NMP′ and P′MN′ are together equal to a right angle, sin P′MN′ = cos NMP′. Hence, since the angle P′MN′ is equal to α,

$$\sin \alpha = \sin P'MN' = \cos NMP' = \frac{\overline{NM}}{P'M}.$$

Also, since PN and N′M are parallel, the angles PMN′ and NPM are equal, and hence

$$\sin \beta = \frac{\overline{NM}}{PM}.$$

Substituting these values of sin α and sin β, we get

$$\mu = \frac{\overline{NM}}{P'M} \cdot \frac{PM}{\overline{NM}} = \frac{\overline{PM}}{P'M}.$$

Now if the point M is taken very near N, $\overline{PM}$ will be very nearly equal to $\overline{PN}$, and $\overline{P'M}$ to $\overline{P'N}$. Hence for rays incident near N we get

$$\mu=\frac{\overline{PN}}{\overline{P'N}},$$

or

$$P'N=\frac{\overline{PN}}{\mu},$$

which gives the distance of the image from the surface of separation between the media. As we have supposed that the medium below AB is denser than that above, and as μ was the refractive index from the upper to the lower medium, μ must be greater than unity, so that the image P′ is nearer to the surface than the object. In the case of water, it is a matter of everyday observation that objects in the water appear nearer the surface than they are really.

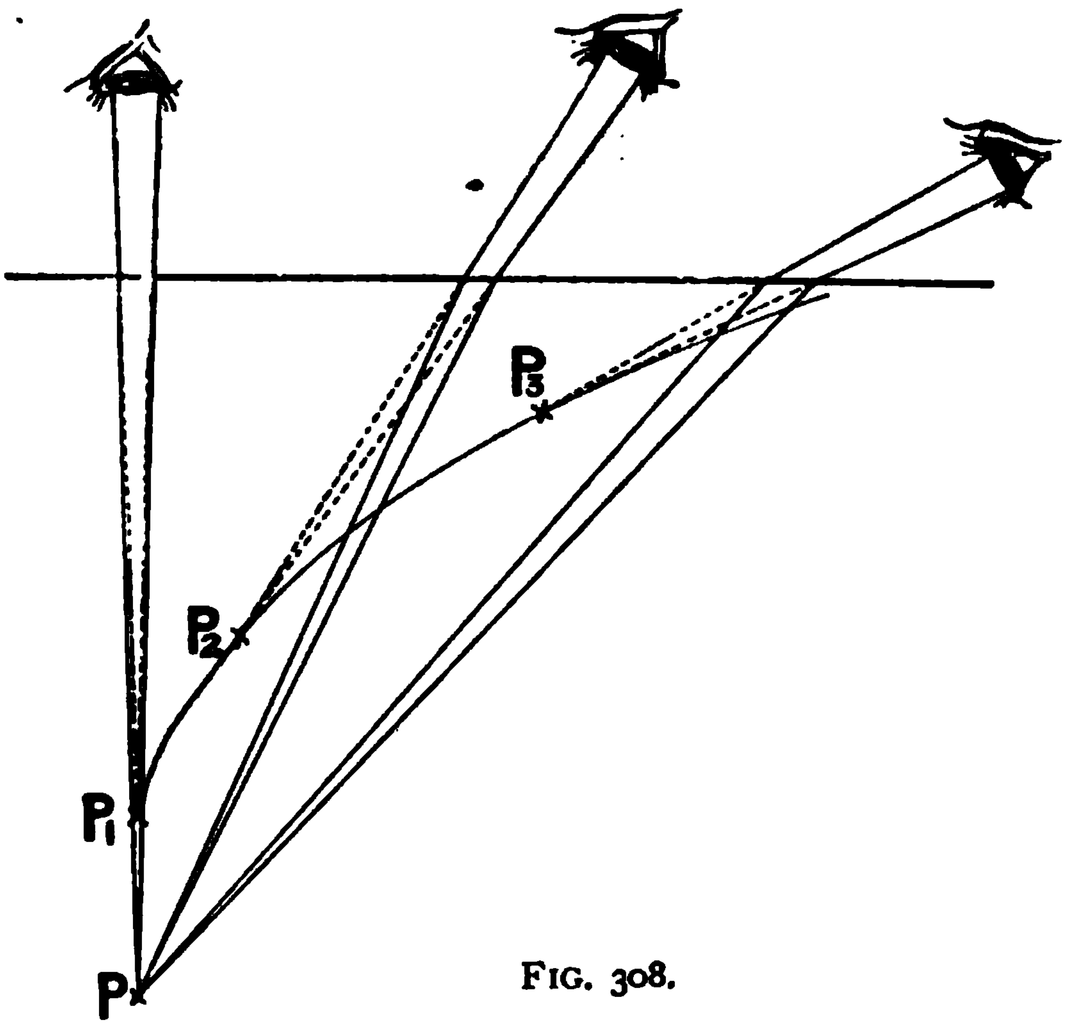


FIG. 308.

For any other rays, except those which are incident almost normally, the directions of a *small* pencil of rays, such as would enter the eye after refraction, very nearly pass through a single point which lies on the caustic, so that an image is formed at this point. The position of the image differs, however, with the position of the eye, as is shown in Fig. 308.

344. Total Internal Reflection.—The equation expressing Snell's law may be written

$$\sin\beta=\frac{\sin\alpha}{\mu}.$$

If μ is greater than unity, *i.e.* if we are considering a ray travelling from

a less dense to a more dense medium, since sin α cannot be greater than unity, the quotient sin α/μ must always be less than unity. Hence for any value of α we can get a corresponding value for β. That is, whatever the value of the angle of coincidence (of course this angle must be less than 90°), there will be a refracted ray.

If, however, we are considering a ray passing from a more dense medium to a less dense one, so that μ is less than unity, then if sin α is less than μ, the quotient (sin α)/μ will be greater than unity. Now sin β cannot be greater than unity, so that we cannot obtain a value for the angle of refraction. When sin α is less than μ, we can obtain a value of β, and there is a refracted ray. When sin α is equal to μ, the quotient sin α/μ is unity, and therefore sin $\beta = 1$, *i.e.* $\beta = 90°$. This means that for this angle of incidence the angle of refraction is 90°, and hence the refracted ray in the less dense medium is parallel to the surface of separation between the media, and just grazes this surface. For larger angles of incidence there is no refracted ray, so that none of the light passes out of the denser medium, it all being reflected back at the surface of separation according to the ordinary laws of reflection. In this case the ray is said to suffer total internal reflection. The angle of incidence, of which the sine is equal to the refractive index, is called the *critical angle.*

If the critical angle (δ) between two media is measured, we can obtain the refractive index from the relation

$$\mu = \frac{1}{\sin \delta}.$$

In the case of glass and air the refractive index is 1.5, and hence the critical angle is given by

$$\sin \delta = \frac{1}{\mu} = .67\,..$$

$$\therefore \qquad \delta = 42°.$$

If a ray of light is incident normally on one of the shorter faces AC of a right-angled glass prism ABC (Fig. 309), it will enter the glass without refraction, and will be incident on the hypotenuse at an angle of 45°. As this angle is greater than the critical angle, the ray will not be able to pass out into the air, but will be totally reflected along O′O″, and will be incident normally at O″ and continue along O″R. The prism has therefore acted as a plane mirror and simply reflected the ray, turning its direction through a right angle.

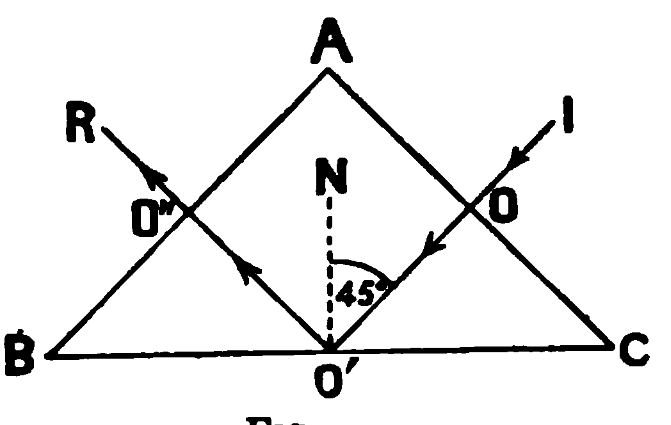


FIG. 309.

345. Refraction through a Prism.—A portion of a refracting medium bounded by two plane surfaces which are inclined at a finite angle is called a prism.

The two plane surfaces are called the faces of the prism, the line in which the faces meet, or would meet if produced, is called the edge of the prism, and the angle between the faces is called the refracting angle, or simply the angle of the prism.

Any plane perpendicular to the two faces, and hence also to the edge, is called a principal plane of the prism.

Since a principal plane is perpendicular to both faces, if a ray of light is incident in the principal plane it will continue in this plane both during its passage through the prism and after leaving the prism.

When the medium of which the prism is composed is denser than the surrounding medium, the ray of light incident in a principal plane will be deviated towards the thick end of the prism. The angle through which the ray has been deviated during its passage is called the angle of deviation.

If the angle of incidence of the incident ray with the first face is altered, it is found that for one angle of incidence the angle of deviation

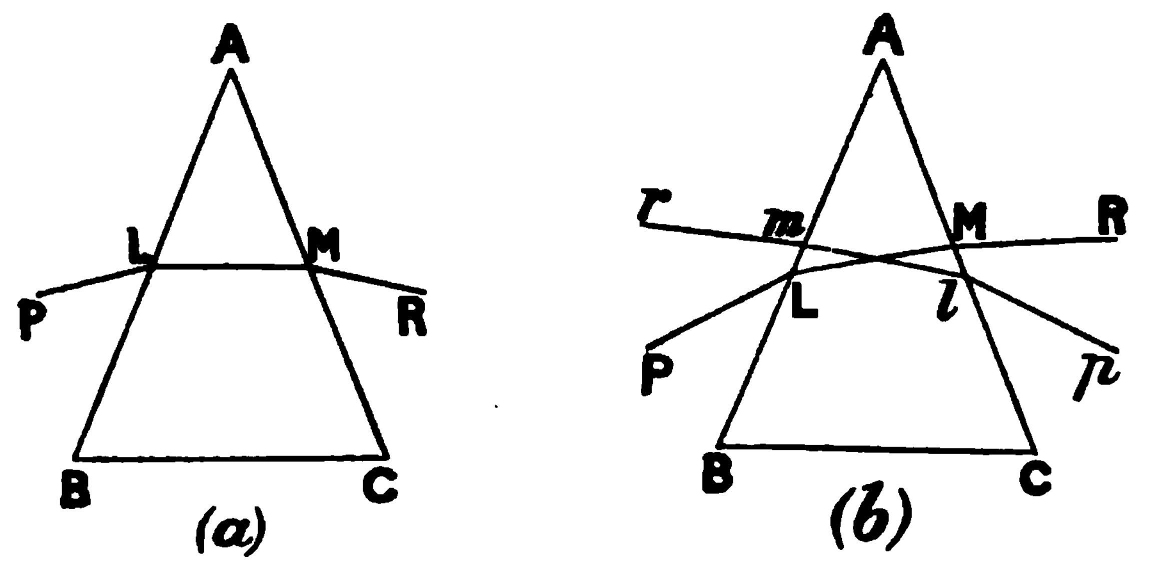


FIG. 310.

produced is a minimum, the deviation being greater both for smaller and larger angles of incidence. The angle through which the ray is deviated under these circumstances is called the angle of minimum deviation.

If a ray of light PL is incident at such an angle on the face AB of the prism ABC, Fig. 310 (*a*), that the deviation is a minimum, the path of the ray in the prism is such that AM is equal to AL. If this were not so, let us suppose that PLMR, Fig. 310 (*b*), represents the path of a ray when the deviation is a minimum. Then a ray of light incident along RM would travel along RMLP, and hence would also suffer minimum deviation, for if we reverse the direction of a ray of light, it always retraces its path.

Next take the point l, such that $\overline{Al}=\overline{AL}$, and m, such that $Am=\overline{AM}$,

and join *lm*. Then draw *pl* inclined at the same angle to AC as is PL to AB, and *mr* inclined to AB at the same angle as is MR to AC. Then the path *plmr* is exactly similar to the path PLMR, and hence a ray incident along *pl* would travel along *plmr*, and would be deviated through the same angle as is the ray PLMR, that is, it would suffer minimum deviation. Hence there are two rays, *pl* and RM, at different angles of incidence, both of which undergo *minimum* deviation, which is impossible, since by experiment there is only one angle of incidence which fulfils this condition. Hence PLMR cannot be at minimum deviation. In the same way, it can be shown that no ray which does not cut the two faces so as to make AL equal to AM can be at minimum deviation.

346. Determination of Refractive Index from the Angle of Minimum Deviation.—From the knowledge of the refracting angle of a prism and the angle of minimum deviation, we can calculate the refractive index from the medium surrounding the prism to the medium of which the prism is composed. Since in practice we have almost always to consider the passage of light from *air* into some other medium, we shall in future refer to the refractive index from air into a medium simply as *the* refractive index of the medium.

Let ABC (Fig. 311) be the trace of a prism, the paper being a principal plane, and PLMR the path of a ray which is at minimum deviation, so that $\overline{AL} = \overline{AM}$. At L and M draw the normals NLN″ and N′MN″, also produce the direction of the emergent ray MR back to D, and produce the direction of the incident ray PL to cut this at E. Then the angle DEL or δ is the angle through which the ray is deviated, and hence δ is by supposition the angle of minimum deviation.

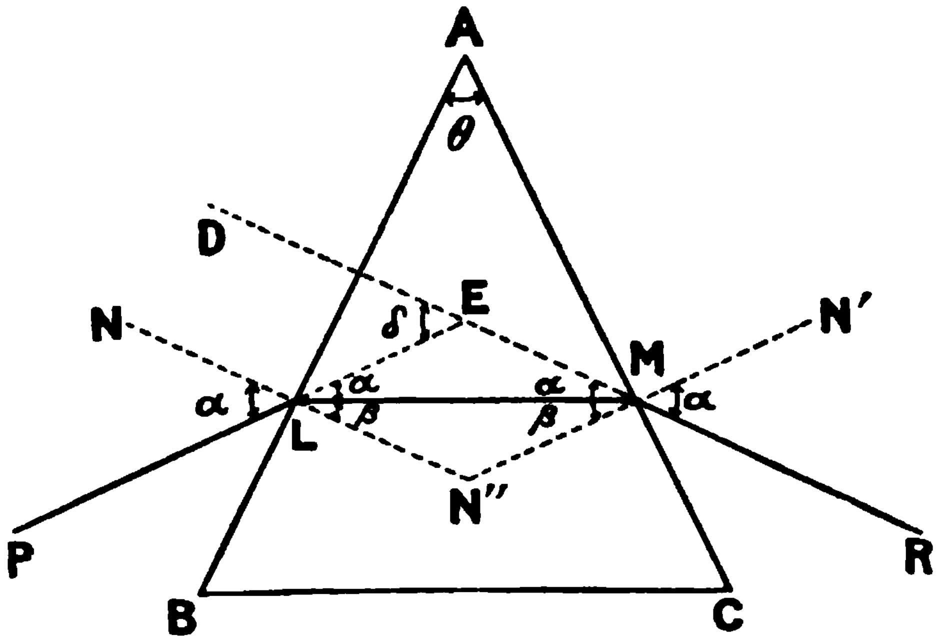


FIG. 311.

Since $\overline{AL} = \overline{AM}$ the angle ALM is equal to the angle AML, and hence as the angles ALN″ and AMN″ are each a right angle, the angle N″LM is equal to the angle N″ML. In the quadrilateral ALN″M, the angles ALN″ and AMN″ are right angles, hence the angle LN″M $= \pi - \theta$. But since LMN″ is a triangle, the angle LN″M $= \pi - 2\beta$. Hence

$$\pi - 2\beta = \pi - \theta$$

$$\therefore \quad \beta = \frac{\theta}{2}.$$

In the triangle ELM the angles ELM and EML are each equal to $\alpha-\beta$, and hence the exterior angle DEL is equal to $2(\alpha-\beta)$, or $\delta=2(\alpha-\beta)$.

$$\therefore \qquad \alpha=\frac{\delta}{2}+\beta=\frac{\delta+\theta}{2}.$$

But if μ is the refractive index,

$$\mu=\frac{\sin\alpha}{\sin\beta}=\frac{\sin\frac{1}{2}(\delta+\theta)}{\sin\frac{1}{2}\theta},$$

and so if the angles θ and δ are measured, the refractive index can at once be calculated.

347. Absolute Refractive Index, and Change in Refractive Index with Change in the Physical Condition of the Medium.— In almost all cases the refractive index from air to a given medium is what we obtain by experiment. When, however, we are making comparisons between the optical properties of different media, it is convenient to eliminate the effect of the medium air, and to consider the refractive index from a vacuum to the medium considered. This is called the absolute refractive index, and can at once be calculated from the refractive index in air, if we know the absolute refractive index of air, by the method given in § 342. Since the absolute refractive index of air under standard condition of pressure and temperature (76 cm. of mercury and 0° C.) is 1.00029, the absolute refractive index for any medium is obtained by multiplying the refractive index relative to air by 1.00029.

According to the electro-magnetic theory of light, to which we shall refer later, the refractive index of a substance is connected with the density d in such a way that the expression

$$\frac{1}{d}\frac{\mu^2-1}{\mu^2+2}=\text{a constant } (R, \text{ say}).$$

This expression is due to Lorentz, and the value of the constant R is independent of the temperature, pressure, and the state of the substance.

Lorentz's formula may be written—

$$\frac{1}{d}\frac{(\mu-1)(\mu+1)}{\mu^2+2}=R.$$

Hence if $\frac{\mu+1}{\mu^2+2}$ is constant, the formula will reduce to $(\mu-1)/d=$constant. From the results of their experiments Gladstone and Dale had previously found that the expression $(\mu-1)/d$ remained constant when the temperature and pressure changed.

In the case of gases the refractive index is always very nearly unity, so that the quantities $\mu+1$ and μ^2+2 are very nearly 2 and 3 respectively,

so that, for gases, Lorentz's expression reduces to that of Gladstone and Dale.

In the following table the values of *R* are given for some substances both in the liquid and gaseous state :—

	$\frac{1}{d}\frac{\mu^2-1}{\mu^2+2}$.	
	Liquid.	Vapour.
Water	0.2061	0.2068
Carbon bisulphide	0.2805	0.2898
Chloroform	0.1790	0.1796

CHAPTER III

LENSES—MEASUREMENT OF REFRACTIVE INDEX

348. Lenses.—A portion of a refracting medium, bounded by two surfaces, one of which is spherical and the other is plane or spherical, is called a *lens.*

If the two surfaces of the lens are spherical, the line joining the centres of the spheres is called the axis of the lens ; if one of the surfaces is plane, the axis is the line drawn through the centre of the sphere perpendicular to the plane.

If the rays proceeding from a point P on the axis, after refraction at the lens, pass through a point P′ ; or if, although they do not actually pass through this point, their directions pass through P′, then P and P′ are called conjugate foci.

The point through which the refracted rays, or their direction, pass when the incident rays are parallel to the axis, is called the *principal focus* of the lens, and the distance between this point and the lens is called the focal length of the lens. Every lens has two principal foci, one on either side of the lens, and at equal distances from the lens.

If AB (Fig. 312) represents a section of a lens, and XX′ the axis, the surfaces of the lens at M and N will be normal to the axis, and hence a ray of light incident along the axis will not be deviated by its passage through the lens. Also if P and Q are any two points, such that the tangents to the lens at P and Q are parallel, a ray of light incident at P in such a direction that it travels along PQ will, after it leaves the lens, travel along QR in a direction parallel to its original path, IP ; for as far as this ray is concerned the lens acts simply as a parallel-sided slab. The point O, where this ray cuts the axis, is called the optical centre of the lens, and is such that all rays which pass through it are

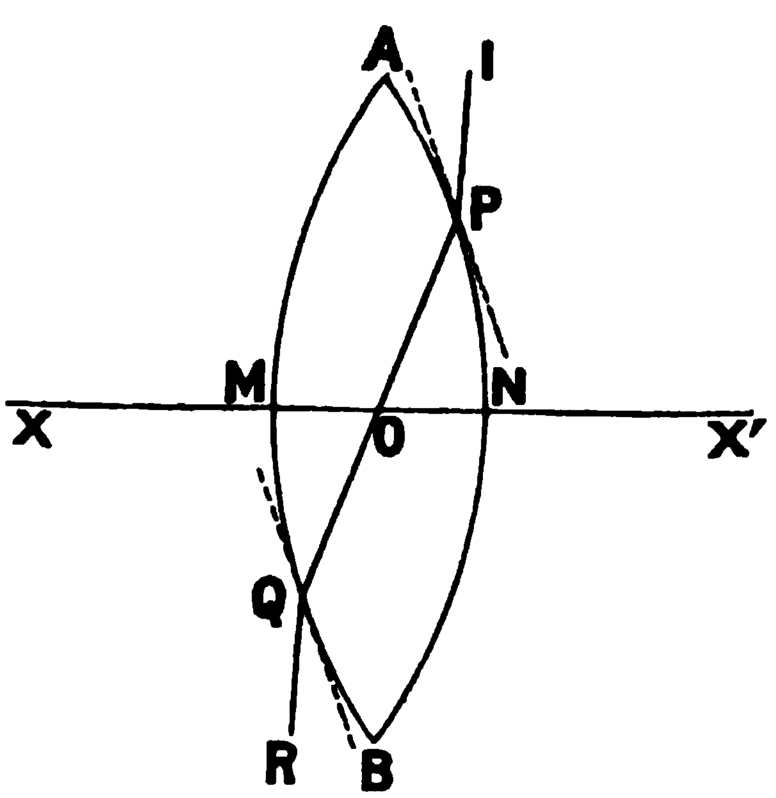


FIG. 312.

undeviated by their passage through the lens. The position of the optical centre varies with the curvature of the surfaces, and may lie quite outside the lens.

In the case of most lenses used in practice the thickness of the lens is small compared with the focal length, so that the points M, O, and N are near together. Such a lens is called a thin lens, and we shall, unless it is specially mentioned, restrict ourselves to thin lenses, so that we may take either of the three points M, O, or N, as the optical centre of the lens.

Lenses are divided into two classes. The first class, called convex lenses or converging lenses, are such that when a pencil of rays parallel to the axis passes through the lens, they are refracted so as to pass through the principal focus. The second class, called concave or diverging lenses, are such that when a pencil of rays parallel to the axis passes through the lens they are refracted, so that although they do not actually pass through the principal focus, yet their directions pass through the focus.

It is only when the medium of a lens is, as is generally the case, denser than the surrounding medium that the above definitions hold. In the opposite case a concave lens is a converging lens, and *vice versa*.

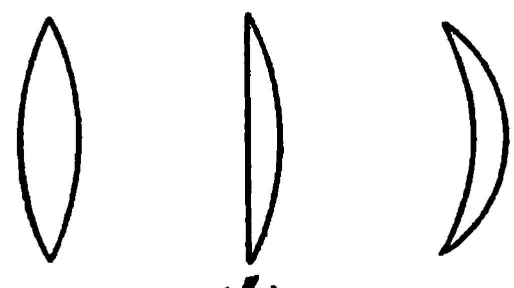


FIG. 313.

In Fig. 313 are given the sections by a plane containing the axis of the three typical forms of convex lens. Lens (*a*) is called a double convex lens, (*b*) a plano-convex lens, and (*c*) a convexo-concave lens or convex meniscus. Lens (*c*) has one convex and one concave surface; the radius of the convex surface is, however, less than that of the concave. In Fig. 314 the three typical forms of concave lenses are shown. Lens (*a*) is called a double concave lens, (*b*) a plano-concave, and (*c*) a concavo-convex or concave meniscus. In (*c*) the concave surface has a smaller radius of curvature than the convex surface.

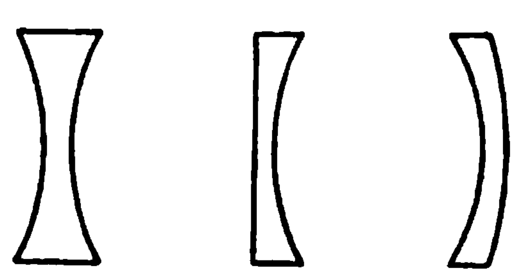


FIG. 314.

If a pencil of rays, all parallel to the axis, falls on a convex lens, such as AB (Fig. 315), after refraction through the lenses they all pass through the principal focus F, and $\overline{\text{OF}}$ is the focal length of the lens. Since the focal length, measured *from* the lens, is in the *same* direction as that in which the incident light is proceeding, it is negative, the same convention as to sign being adopted as in the case of mirrors (§ 337).

If a pencil of rays parallel to the axis falls on a concave lens, such as CD (Fig. 315), after their passage through the lens the rays diverge and travel as if they came from the principal focus F′, the point F′ being on

the side of the lens on which the light is incident. Hence as $\overline{OF'}$ is measured in the opposite direction to the incident light it is positive, so that the focal length of a concave lens is positive. It is for the above reasons that convex lenses are sometimes called negative lenses, while concave lenses are called positive.

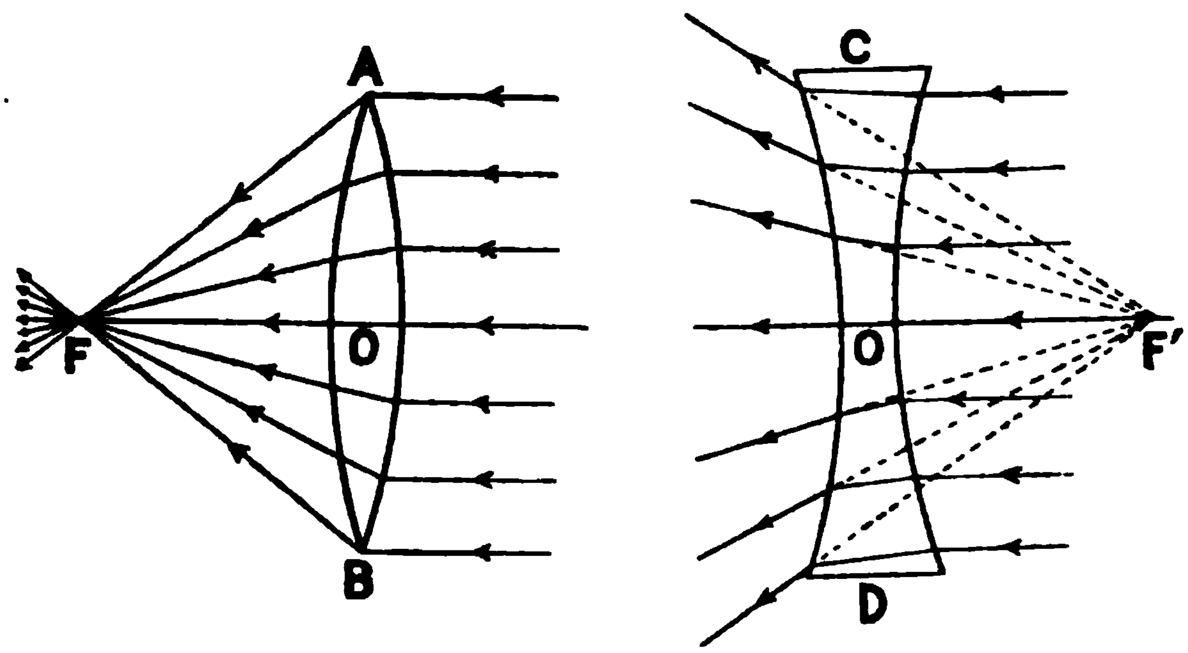


FIG. 315.

If r_1 is the radius of curvature of the first surface on which a parallel beam of light is incident, r_2 the radius of curvature of the other surface, f the focal length, and μ the refractive index of the medium of which the lens is composed, then these quantities, due regard being paid to their proper sign, are connected by the equation—

$$\frac{1}{f} = (\mu - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) \quad . \quad . \quad . \quad (1).$$

This equation can be deduced in the following manner. Let AB (Fig. 316) be a section of a lens, and suppose a pencil of parallel rays falls on the lens parallel to the axis, OF, and is brought to a focus at the principal focus, F. We will consider two rays, I_1AF and I_2OF. If through C we draw CC′ perpendicular to the incident light, CC′ will be a section of the wave-front, for we are dealing with plane waves. Hence the wave will reach the points C and C′ in the same phase. After leaving

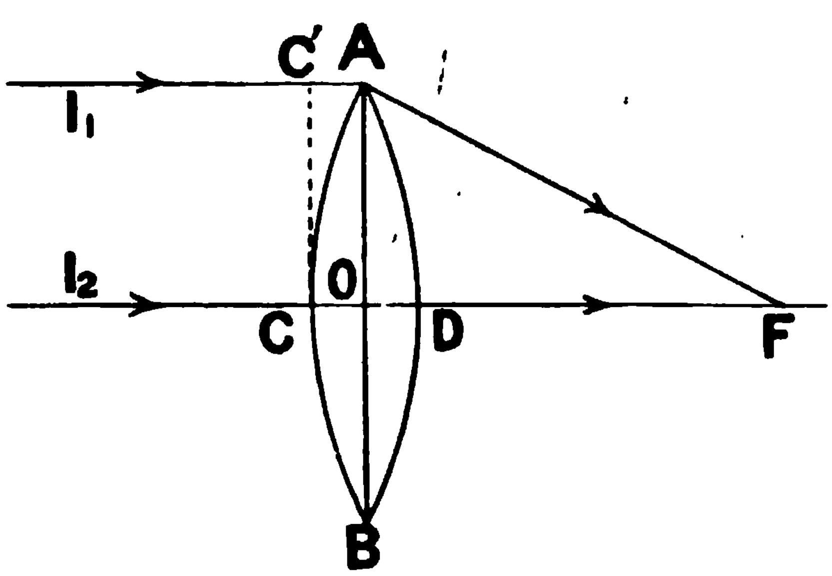


FIG. 316.

C′ the wave will travel along C′A, then, as A is just at the edge of the lens, it will pass through a very small thickness of glass, and then travel

along AF. The part of the wave near C will in the same way travel through a thickness CD of glass, and then over a distance $\overline{DF}$ of air. Now as the rays are brought to a focus at F, the different parts of each wave, such as CC′, must all reach F at the same time, for, as we have seen in the case of a real focus, the wave-front is a sphere with positive curvature, and the wave-fronts eventually are reduced to points at the focus. Thus the time taken by the one portion of the wave to travel over the path C′AF must be the same as that taken by the other portion to travel over the path CDF.

We shall see later (§ 366) that the velocity of light in a medium, of which the refractive index is μ, is $1/\mu$ of the velocity of light in air. Hence it will take the light the same time to travel over a distance μx in air, as it does over a distance x in a medium of refractive index μ.

If we call the distance AO a, the radius of curvature of the surface ACB r_1, that of the surface ADB r_2, and the focal length, $\overline{OF}$, of the lens f, then, from the well-known property of the segments of chords of a circle, we have

$$\overline{CO}\,(2r_1 - \overline{CO}) = a^2,$$

or

$$\overline{CO} = a^2/2r_1,$$

for if the lens is thin, and it is only for thin lenses of small curvature that our investigation holds, we may neglect the term $\overline{CO}^2$. In the same way

$$\overline{OD} = a^2/2r_2.$$

Now the length of the path C′AF is

$$\begin{aligned}\overline{C'A} + \overline{AF} &= CO + \sqrt{(\overline{OF}^2 + \overline{AO}^2)}\\ &= a^2/2r_1 + \sqrt{(f^2 + a^2)}\\ &= a^2/2r_1 + f\sqrt{(1 + a^2/f^2)}\\ &= a^2/2r_1 + f(1 + a^2/2f^2)\\ &= a^2/2r_1 + f + a^2/2f,\end{aligned}$$

where terms in a^4/f^4 and higher powers have been neglected, for f is much greater than a. Also the length of the path CDF, allowance being made for the slower velocity of the waves in glass, is

$$\mu(\overline{CO} + \overline{OD}) + \overline{DF} = \frac{\mu a^2}{2}\left(\frac{1}{r_1} + \frac{1}{r_2}\right) + f - \frac{a^2}{2r_2}.$$

Hence if the length of the two paths are equal, so that the two portions of the wave reach F at the same instant, we have

$$\frac{a^2}{2r_1} + f + \frac{a^2}{2f} = \frac{\mu a^2}{2}\left(\frac{1}{r_1} + \frac{1}{r_2}\right) + f - \frac{a^2}{2r_2},$$

or

$$\frac{a^2}{2f} = \frac{\mu a^2}{2}\left(\frac{1}{r_1} + \frac{1}{r_2}\right) - \frac{a^2}{2}\left(\frac{1}{r_1} + \frac{1}{r_2}\right),$$

$$\frac{1}{f} = (\mu - 1)\left(\frac{1}{r_1} + \frac{1}{r_2}\right) \quad . \quad . \quad . \quad (1).$$

If one of the faces, say the second, is plane, so that $r_2=\infty$, the above formula will become

$$\frac{1}{f}=\left(\frac{\mu-1}{r_1}\right) \quad . \quad . \quad . \quad (2),$$

since when $r_2=\infty$, $\frac{1}{r_2}$ will be zero.

As an illustration of the application of these formulæ, suppose that the radius of curvature of the right-hand surface of the lens AB (Fig. 316) is 10 cm., and that of the left-hand surface is 8 cm., and we required to find the focal length, the refractive index of the medium (glass) of which the lens is composed being 1.33. In this case the centre of the sphere of which the right-hand face is a part is to the left of the lens, and hence its radius of curvature measured from the lens is measured to the left, *i.e.* in the same direction as the incident light, and is therefore negative. In the same way, the radius of curvature of the other face is positive. Hence in the example $r_1=-10$ cm., and $r_2=8$ cm. Putting these values into equation (1), we get

$$\begin{aligned}\frac{1}{f}&=(1.33-1)\left(-\frac{1}{10}-\frac{1}{8}\right)\\&=.33(-.100-.125)\\&=-0.0742.\\ \therefore \quad f&=-13.48 \text{ cm.}\end{aligned}$$

Next suppose that the radius of curvature of the right-hand surface of a double concave lens CD (Fig. 315) is 10 cm., and that of the left surface is 8 cm. In this case $r_1=+10$ cm. and $r_2=-8$ cm. Hence

$$\begin{aligned}\frac{1}{f}&=(1.33-1)\left(\frac{1}{10}+\frac{1}{8}\right)\\&=.33(.100+.125)\\&=0.0742.\\ \therefore \quad f&=13.48 \text{ cm.}\end{aligned}$$

The relative distances of the image and object from a lens are given by the formula

$$\frac{1}{f}=\frac{1}{v}-\frac{1}{u} \quad . \quad . \quad . \quad (3),$$

where f is the focal length of the lens and u and v are the distances of the image and object respectively from the lens, each of these quantities being taken with its appropriate sign.

349. Geometrical Construction for finding the Image and Relative Size of Image and Object.—The position of the image formed by a lens can be found by a geometrical construction exactly similar to that used in regard to mirrors in § 338.

We take one ray which, proceeding from the point Q (Fig. 317) of the object, passes through the centre of the lens O, and hence is undeviated, and another, QM, which is parallel to the axis, and hence after passing through the lens either actually passes, or its direction passes through the principal focus F. The point Q′ where the two rays intersect is the image of Q. If the rays actually pass through Q′ the image is real; if the rays do not actually pass through Q′, but only their directions when

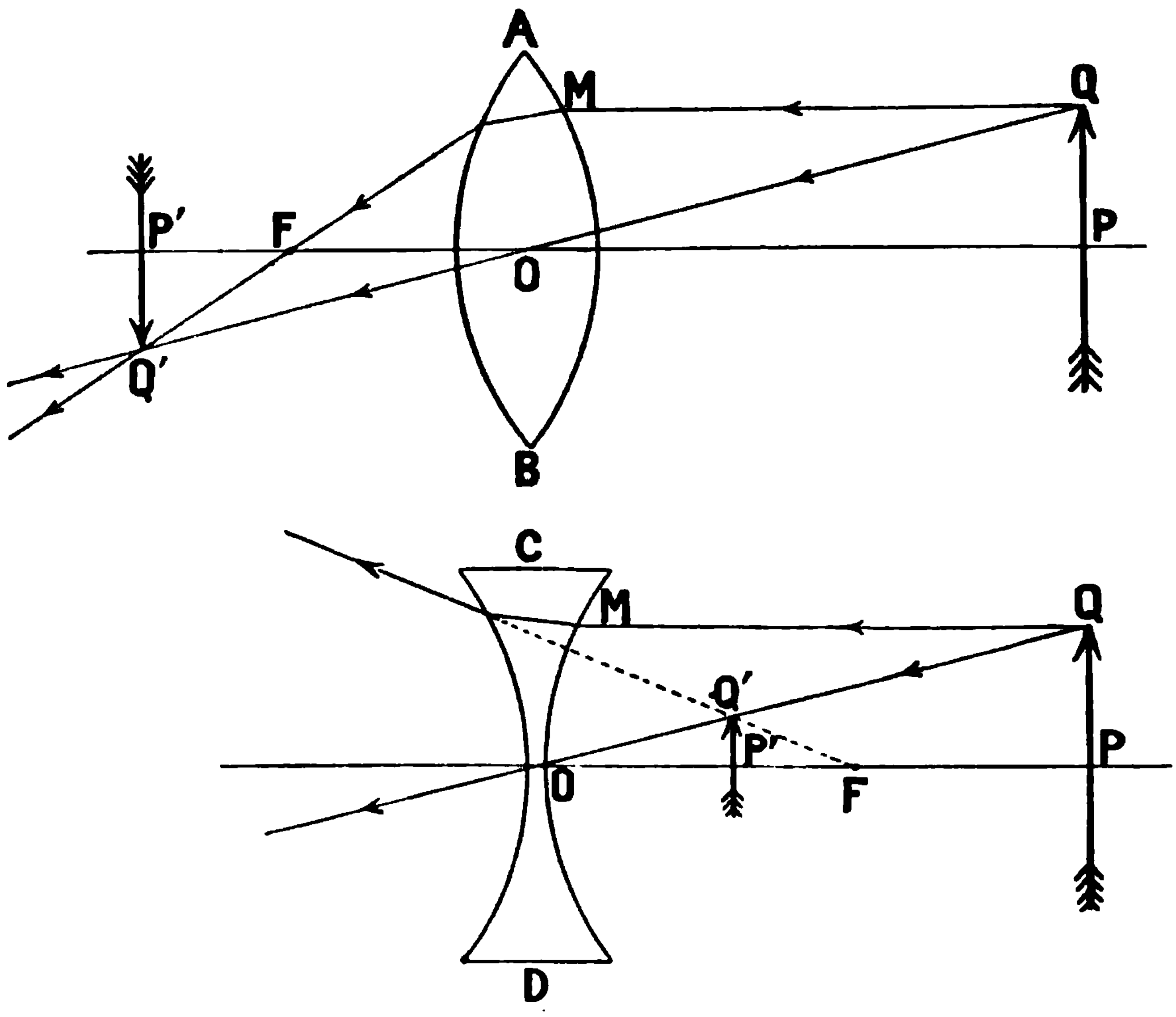


FIG. 317.

produced backwards, then the image is virtual. In Fig. 317 the image produced by the lens AB is real, while that produced by the lens CD is virtual.

In the figure the triangles QOP and Q′OP′ are similar. Hence

$$\frac{QP}{Q'P'}=\frac{OP}{OP'}=\frac{u}{v}.$$

Thus the size of the object is to the size of the image as the distance of the object from the lens is to the distance of the image from the lens.

The positions, &c., of the image for different positions of the object

are given in the two following tables, and the results can easily be verified by drawing figures in the different cases :—

CONVEX LENSES.

Distance of Object from Lens.	Distance of Image from Lens.	Character of Image.		
∞	f	...	...	...
Between ∞ & $2f$	Between $-f$ & $-2f$	Real	Inverted	Diminished
Between $2f$ & f	Between $-2f$ & $-\infty$	Real	Inverted	Magnified
Between f & 0	Between $+\infty$ & 0	Virtual	Erect	Magnified

CONCAVE LENSES.

Distance of Object from Lens.	Distance of Image from Lens.	Character of Image.		
∞	f	...	...	...
Between ∞ & 0	Between $+f$ & 0	Virtual	Erect	Diminished

350. Position of the Image formed by Two Lenses.—Suppose two lenses AB and CD (Fig. 318) to be placed so that their axes coincide, and at a distance d apart, the focal length of AB being f_1, and that of CD f_2, and that we require to find the position of the image of a luminous point P at a distance u from the first lens formed by the combination of the two lenses.

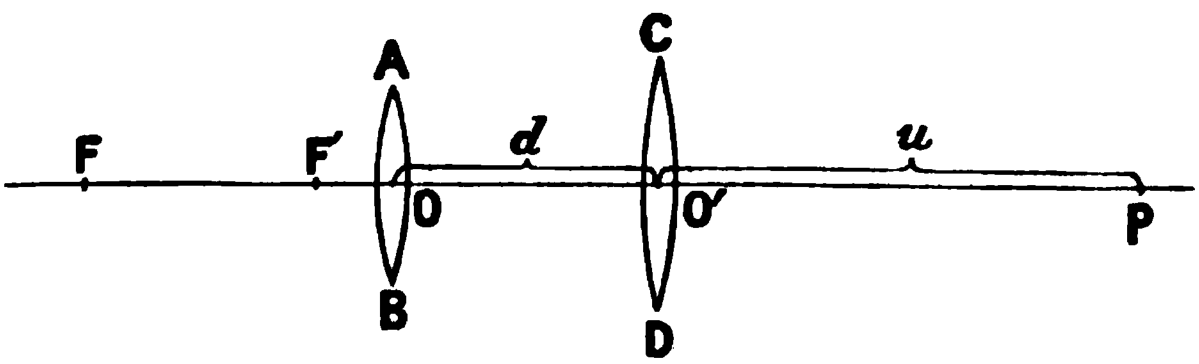


FIG. 318.

If the lens CD existed alone, then the distance v of the image from O′ will be given by

$$\frac{1}{v}=\frac{1}{f}+\frac{1}{u} \quad . \quad . \quad . \quad (1).$$

The distance of this image from the second lens is $v+d$, since if v is positive, *i.e.* the image is to the right of CD, the distance from O is equal to the distance from O′ plus d, and with reference to O this length would be positive. If the image formed by the first lens is to the left of CD, v is negative, and $v+d$ still gives the distance of the image from O, with its proper sign. Hence, treating this image as the object for the second lens, if v' is the distance of the image of this image from O, we have

$$\frac{1}{v'}=\frac{1}{f_2}+\frac{1}{v+d} \quad . \quad . \quad . \quad (2).$$

By substituting the value of v given by (1) in this equation the value of v' can be obtained, which gives the position of the image formed by the combination.

If the two lenses are placed in contact and their focal lengths are sufficiently large, we may take $d=0$. Under these conditions (2) reduces to

$$\frac{1}{v'}=\frac{1}{f_2}+\frac{1}{v}.$$

Hence, substituting from (1) the value of $\frac{1}{v}$, we get

$$\frac{1}{v'}=\frac{1}{f_2}+\frac{1}{f_1}+\frac{1}{u},$$

or

$$\frac{1}{f_1}+\frac{1}{f_2}=\frac{1}{v'}-\frac{1}{u} \quad . \quad . \quad . \quad (3).$$

If the object is at an infinite distance ($u=\infty$) the incident light is a parallel beam, and the light is brought to a focus at a point at a distance F from the two lenses, where F is the value of v' obtained when $u=\infty$.

Thus

$$\frac{1}{F}=\frac{1}{f_1}+\frac{1}{f_2} \quad . \quad . \quad . \quad (4).$$

The quantity F may be called the focal length of the combination, and the above equation shows that the reciprocal of the focal length of a combination of two lenses in contact is equal to the sum of the reciprocals of their focal lengths.

351. The Eye.—The eye consists practically of a system of lenses by means of which a real image of external objects is formed on a network of nerves, called the retina, at the back of the eye, which nerves convey the impression of vision to the brain.

A diagrammatic section of the eye is shown in Fig. 319. The eye is surrounded, except in front, by a horny opaque coat, the sclerotic. The front transparent portion of this outside coating is called the cornea, C. The inside of the eye is divided into two portions by the iris I, the crystalline lens L, and the muscles which attach the latter to the walls of the eye. The crystalline lens is a double convex lens, of which the anterior surface has a radius of curvature of about 1.1 cm., while the posterior surface has a radius of curvature of about 0.8 cm. By means of the muscles attached to the edge of the lens the curvature of the faces, and hence the focal length, can be altered at will. The iris I forms an opaque coloured diaphragm perforated by a central opening called the pupil. The dia-

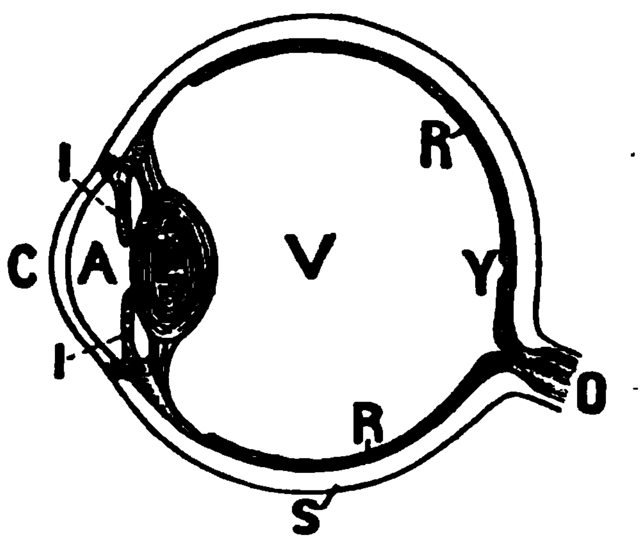


FIG. 319.

meter of the pupil varies with the intensity of the light which enters the eye; thus in a strong light the pupil is contracted, while in a feeble light it is expanded, these movements being involuntary. The space between the cornea and the lens is filled by a transparent liquid called the aqueous humour, while that between the lens and the retina R is filled with a liquid called the vitreous humour.

The retina consists of a semi-transparent network of nerve fibres formed by the spreading out of the termination of the optic nerve. Near the centre of the retina there is a round yellowish spot Y, called the yellow spot, and vision is most distinct when the image falls on this spot. The point of the retina where the optic nerve enters is insensitive to light, so that when the image of an object falls on this spot, which is called the blind spot, no sense of vision is produced.

If the eye really consisted of an ordinary lens, it is evident that for only one distance would the light from a luminous point be brought to a focus on the retina. At all other distances an indistinct and blurred image would be produced. The eye, however, possesses the power of accommodation so that the images of objects at very different distances can all be formed on the retina. In most cases the eye, when at rest, is so arranged that the image of a distant object is in focus on the retina. The accommodation for nearer objects is produced by a slight forward motion of the lens and an increase of the curvature of its surfaces, the increase in curvature of the front surface being very much the more strongly marked of the two.

The range of the accommodation is not unlimited, so that objects which are very near the eye cannot be clearly seen. Since when an object is a great way off we cannot make out small details about it, and neither can we do so when it is very near, it follows there must be some distance at which we are able to see most distinctly. This distance, which is called the *distance of distinct vision*, is for a normal eye between 25 and 30 cm., or 10 and 12 inches.

352. Defects of Vision.—There are three defects of the eye which are of comparatively frequent occurrence. These are known as (1) short-sight, (2) long-sight, (3) astigmatism.

In the case of short-sight distant objects cannot be seen distinctly, because the point to which the rays from distant objects are brought to a focus is, even when the lens is at its flattest, in front of the retina. We may here consider that the lens is too convergent for the size of the eye-ball, so that if in front of the eye we place a concave lens so as to make, with the lens of the eye itself, a less convergent system than the crystalline lens itself, the defect of short-sightedness can be corrected.

Since the image formed by a concave lens is always virtual, it is evident that if d is the *maximum* distance at which a short-sighted person can see distinctly, then if the concave lens is such that the focus of parallel rays is at a distance d from the eye, the eye will be able to

see clearly this image, and hence all distant objects. Since the spectacle lens is always placed quite close to the eye, the distance of the focus for parallel rays from the lens must be d, that is, the lens must have a focal length d.

In long-sight, or hypermetropia, near objects cannot be seen distinctly, this being due to the fact that, when the lens is as much curved as possible, the image of objects even some distance off is formed behind the retina. In this case the eye, when relaxed, is in such a state that parallel rays meet behind the retina, so that to see distant objects the eye has to be accommodated. This defect can be remedied by placing a convex lens in front of the eye, for by this means the focus of the combination of lens and eye is nearer the crystalline lens than when no spectacle lens is used.

Let the minimum distance at which a long-sighted eye can see clearly be d, and it be required to find the focal length of a convex lens which will produce distinct vision at the ordinary distance of most distinct vision, say D. Then, assuming that the spectacle lens and eye are close together, we must take the focal length f of the lens such that the image produced by an object at a distance D must be on the *same* side of the lens as the object, and at a distance from the lens d, where $d > D$. Here $u = +D$ and $v = +d$. Hence

$$\frac{1}{f} = \frac{1}{d} - \frac{1}{D}.$$

Since d is $> D$, $\frac{1}{d}$ will be less than $\frac{1}{D}$, and hence $\frac{1}{f}$, and therefore f will be negative, and the lens must be convex, which agrees with the conclusion at which we have already arrived.

In astigmatism the surfaces of the cornea, and the lens, but principally the former, are not symmetrical about the axis. In most cases the vertical section of the cornea of an astigmatic eye is more curved than a horizontal section, so that the image of a horizontal line is formed nearer the crystalline lens than the image of a vertical line. This defect is remedied by the use of spectacles in which the surfaces of the lenses are not spheres, but differ from these in the opposite sense to that of the defective eye.

353. The Simple Microscope, or Magnifying Glass.—We have seen in the preceding section that if we attempt to increase the distinctness with which an object can be seen by bringing it nearer the eye, so that it appears larger, a position is at length reached such that if we bring it nearer we are unable to see it at all distinctly.

If a convex lens is placed in such a position that the object AB (Fig. 320) is between the principal focus F and the lens, the rays, after they leave the lens, will proceed as if they came from the virtual image A′B′. This image is found by the construction given in § 349, and is erect

and magnified. Now, although the image A′B′ and the object AB subtend nearly the same angle at the eye, yet if A′B′ is at the distance of distinct vision, on removing the lens the object AB would be seen very indistinctly, it being so much within the minimum distance of distinct vision. Hence, to find the magnification produced by the lens, we must compare the angle subtended at the eye by the image A′B′ with the angle subtended by the object when it is at the minimum distance at which it can be clearly seen, *i.e.* at *ab*. If the eye is very near the convex lens, the angles subtended by the image, and by the object *when at the distance of distinct vision*, are very nearly equal to A′OB′ and *aOb* respectively.

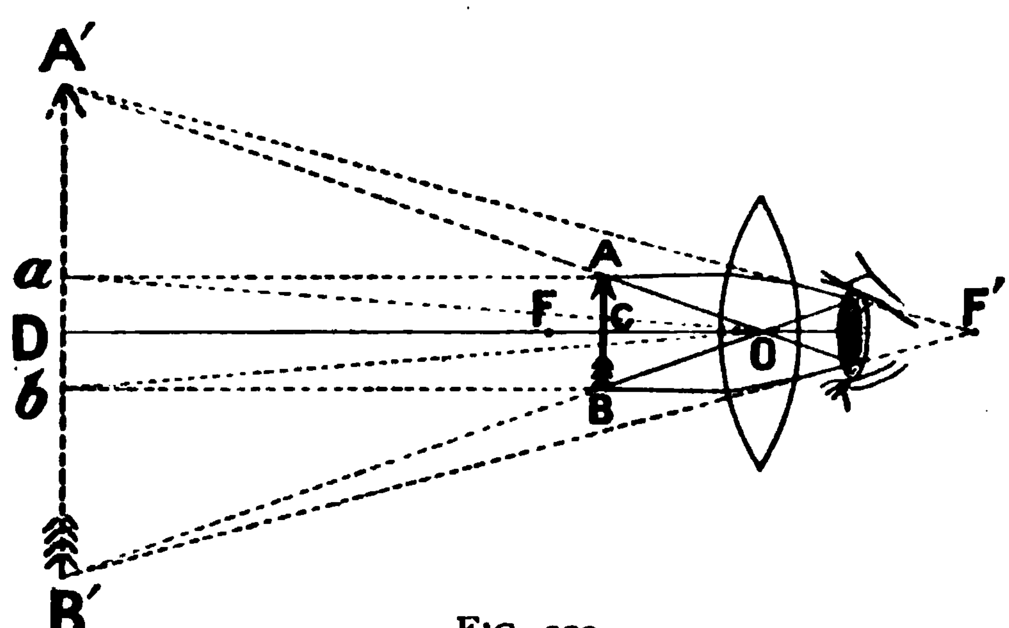


FIG. 320.

Hence the magnification is equal to $\frac{\overline{A'B'}}{\overline{ab}}$ or $\frac{\overline{A'B'}}{\overline{AB}}$, or, since the triangles A′OD, AOC are similar, to $\frac{\overline{DO}}{\overline{CO}}$ But $\overline{CO}$ is the distance of the object from the lens, and $\overline{DO}$ is the distance of the image from the lens, hence we have the magnification $=\frac{v}{u}$. If f is the focal length of the lens,

$$\frac{1}{f}=\frac{1}{v}-\frac{1}{u},$$

or

$$\frac{v}{f}=1-\frac{v}{u},$$

or

$$\frac{v}{u}=1-\frac{v}{f}.$$

Hence the magnification $=1-\frac{v}{f}$; or, since the image is to be formed at the distance of distinct vision D, the magnification $=1-\frac{D}{f}$, where it must be remembered that f is negative and D is positive.

From the above expression it will be seen that the magnification increases as f decreases, so that to obtain great magnifying power a lens of very short focal length must be taken.

354. The Compound Microscope.—In the simple microscope the greatest magnification which can be obtained is about one hundred-fold,

and in order to obtain greater magnification a combination of convex lenses must be used, called a compound microscope. In its simplest form the compound microscope consists of two convex lenses, A and B (Fig. 321). The lens A, or objective, is of short focal length, and is so placed that the object PQ is just beyond its principal focus, so that a real inverted and slightly magnified image is produced at P′Q′. The second

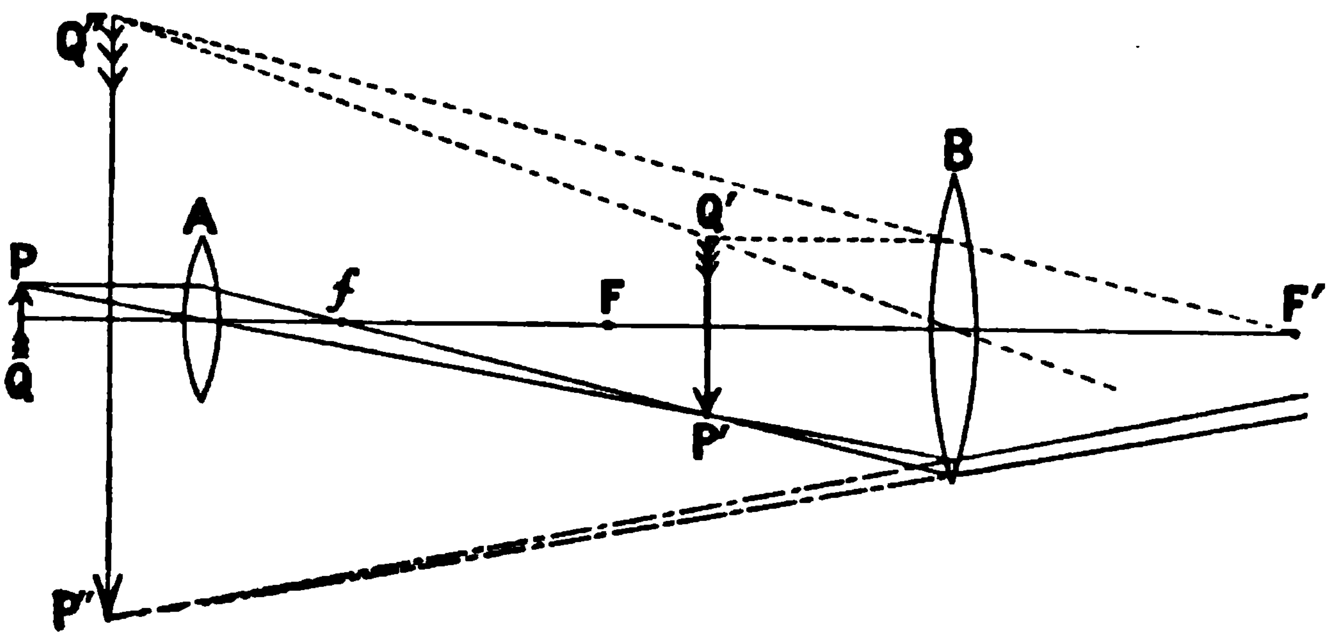


FIG. 321.

lens or eye-piece, B, is placed at such a distance from the objective that the image formed by the latter is just inside the principal focus F, and hence the eye-piece, acting as a simple microscope, gives a virtual and magnified image P″Q″.

Since the normal eye, when at rest, is adjusted for an object at a great distance, *i.e.* for parallel rays, it is less trying to the eyes if, whenever possible, we arrange an optical instrument so that the rays that enter the eye are parallel. In the case of the compound microscope this can be easily done, for if the lens B be placed so that the image P′Q′ is formed at its principal focus, then the rays from each point of the image

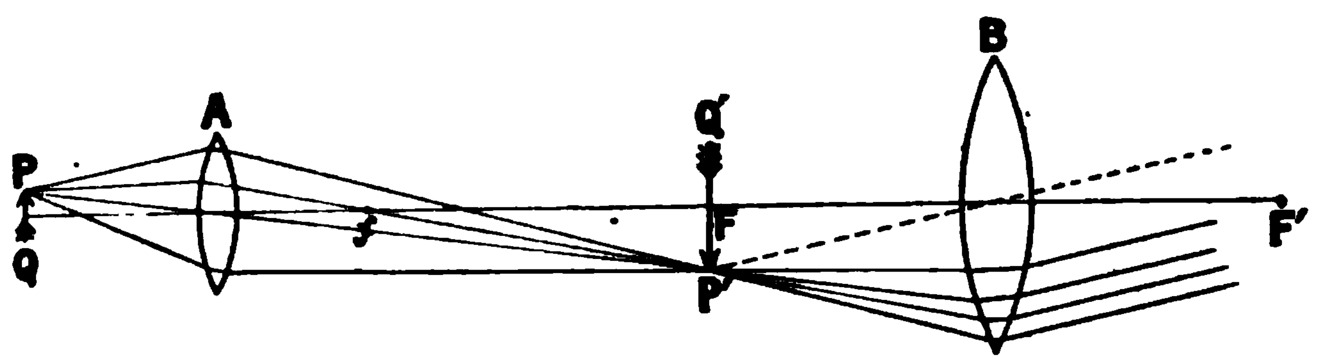


FIG. 322.

P′Q′, after passing through the eye-piece, will emerge parallel, as shown in Fig. 322. The angle between the rays from the extreme points of the object when they enter the eye remains practically the same, as in Fig. 321, so that the apparent size of the image remains the same; we have, however, done the focussing by means of the instrument, instead of using the accommodation of the eye.

The above form of microscope is much simpler than any now used, but to go into the theory of the modern microscope to any purpose would be beyond the scope of this work. On account of spherical aberration and chromatic aberration, both the objective and the eye-piece consist of combinations of lenses, the objectives of some modern high-power microscopes having as many as ten separate lenses.

355. The Telescope.—A telescope is an instrument by which a magnified image of a *distant* object may be produced. If a convex lens is used to form an image of a distant object on a screen the image will be smaller than the object in the ratio of the distance of the object to the distance of the screen from the lens. Suppose the object is at a distance L, so great that the image is practically at the principal focus, *i.e.* at a distance F from the lens. Then if X is the size of the object, and x that of the image, we have—

$$\frac{x}{X}=\frac{F}{L}.$$

If the image is viewed at the distance of distinct vision D, it will subtend an angle which in angular measure may be taken to be $\frac{x}{D}$. Also the object when looked at directly will subtend an angle $\frac{X}{L}$, since by supposition the object is so far off that L practically gives the distance from the eye. Hence the magnification is

$$\frac{x}{D}\Big/\frac{X}{L}=F/D.$$

Thus if the distance of distinct vision is 30 cm., and the focal length of the lens is 60 cm., the magnification produced when a distant object, such as the moon, is viewed will be 2. In other words, the diameter of

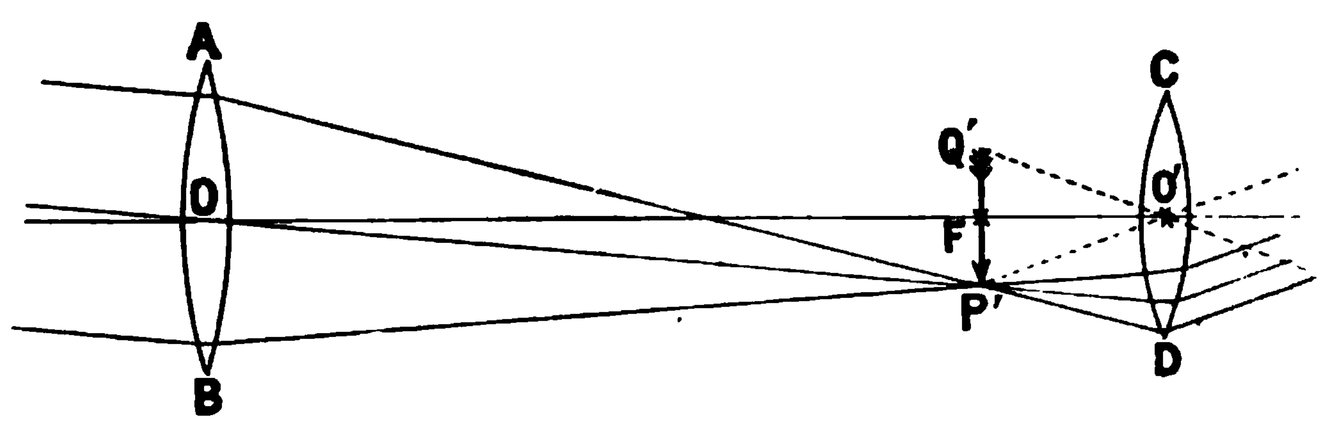


FIG. 323.

the image formed on the retina, when the image on the screen is viewed, will be twice the diameter of the image produced on the retina when the moon is viewed directly.

If instead of receiving the image formed by the convex lens on a screen we view it through an eye-piece lens, then we can obtain further magnifica-

tion. The course of a parallel beam of light through such a combination of lenses, called an astronomical telescope, is shown in Fig. 323. An inverted and diminished image is produced by the object-glass AB at P′Q′, the principal focus. This image, being at the principal focus of the eye-lens, the rays from each point of the object on leaving this lens will be parallel, and entering the eye will produce an image. Since the image P′Q′ is inverted, and the virtual image of this image produced by CD is erect, the final effect is that a magnified inverted image of the object is seen.

The angle made by the rays from the extreme points P′Q′ of the image P′Q′ when they enter the eye is Q′O′P′, which, since this angle is small, is equal to $\frac{Q'P'}{FO'}$ or $\frac{x}{f}$ in circular measure, if f is the focal length of the eye-lens. But

$$\frac{x}{X}=\frac{F}{L},$$

$$\text{or } x=\frac{FX}{L}.$$

Hence the angle between the rays when they enter the eye is

$$\frac{FX}{fL}.$$

But the angle subtended by the object seen without any telescope is $\frac{X}{L}$. Henco the magnification is

$$\frac{FX}{fL}\bigg/\frac{X}{L}=\frac{F}{f}.$$

The magnification thus depends on the ratio of the focal length of the object-glass to that of the eye-piece. Increasing the *diameter* of the object-glass, if the focal length is unaltered, does not change the magnification ; it will, however, increase the brightness of the image, for it will collect more of the rays that leave any particular point of the object and bring them to a point in the image. In practice, both the object-glass and the eye-piece consist of several lenses, so as to avoid chromatic and spherical aberration.

The image seen in an astronomical telescope is inverted, and although this does not matter for astronomical purposes, yet it would be very inconvenient when the telescope is used to view terrestrial objects.

An erect image is obtained by placing two convex lenses between the object-glass and eye-piece. These two lenses are at a distance apart equal to the sum of their focal lengths, and the one nearer the objective is placed so that the image formed by the objective is at the principal focus of this lens. In Fig. 324 P′Q′ is the image formed by the objective, and as this image is at the principal focus of the lens M_1, all the rays

leaving any point of the object will, after their passage through the lens M_1, be parallel. These parallel rays, falling on the lens M_2, form an image $P''Q''$ at the principal focus, F'', of this lens which is inverted with reference to $P'Q'$, and hence erect with reference to the object. This image, $P''Q''$, is viewed by an eye-piece as in the astronomical telescope. The two lenses M_1 and M_2 serve simply to give an erect image, and are fixed at a constant distance apart. The telescope is focussed by altering their distance from the object-glass, so that the image formed by the latter is always at the principal focus of M_1.

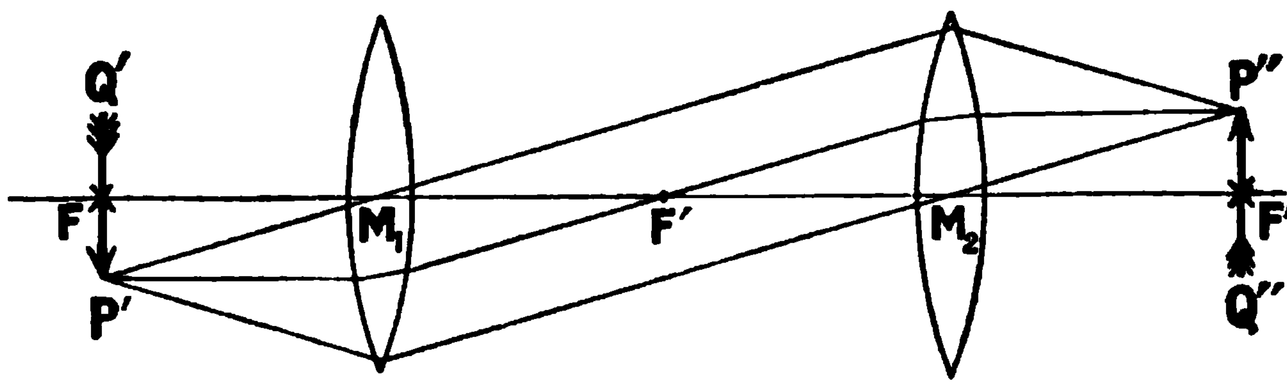


FIG. 324.

A form of telescope which gives an erect image with only two lenses, and which can be made much shorter than the terrestrial telescope described above, is Galileo's. This form of telescope consists of a convex lens as object-glass, AB (Fig. 325), and a concave lens, CD, as eye-piece. If CD were not present the convex lens would form a real image at $P'Q'$; when the concave lens is interposed between AB and the image, so that

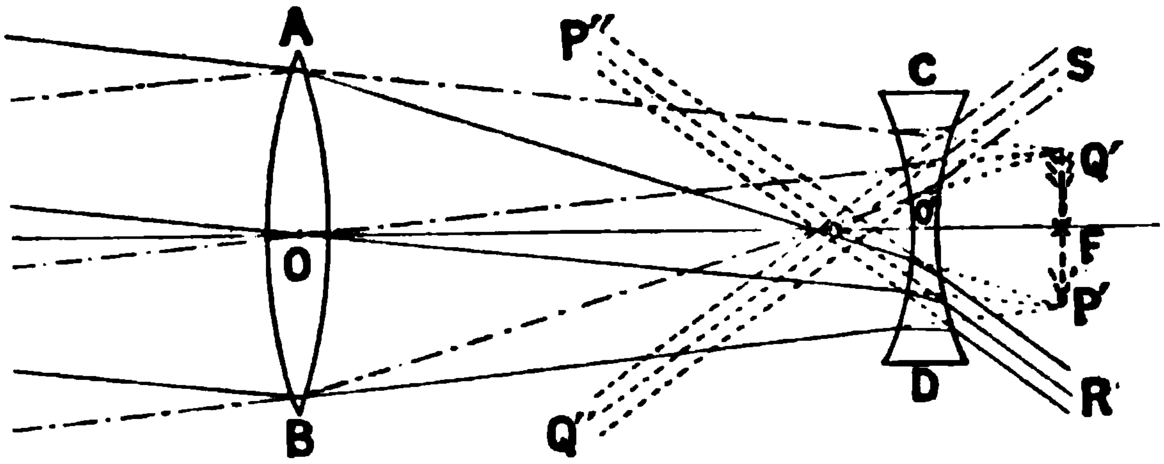


FIG. 325.

the distance $O'F$ is equal to the focal length of CD, the rays of light from any one point of the object will be parallel after they leave this lens. Hence, as shown in the figure in full lines, for a pencil of rays coming from the point P of the object, the rays will enter the eye in the direction $P''R$, while a pencil coming from Q will, as shown by the dotted lines, enter the eye in the direction $Q''S$, so that the eye sees an enlarged and erect image. The magnification is, as before, equal to F/f, where F is the focal length of the objective and f that of the eye-piece. In Galileo's telescope the distance between the objective and eye-piece is $F-f$, while

in the astronomical telescope it is $F+f$, hence the saving in length for an equal magnifying power, and with objectives of equal focal length. Opera and field glasses consist of two Galilean telescopes, one for each eye, the distance between the objectives and eye-pieces being variable by means of a screw, so that the image formed by the objective may always be formed at the principal focus of the eye-piece.

356. The Optical Lantern.—The optical or magic lantern is an arrangement by which an enlarged image of an object can be thrown on a screen. A convex lens or system of lenses is used to form the image, but since the image is considerably magnified, the light which the lens

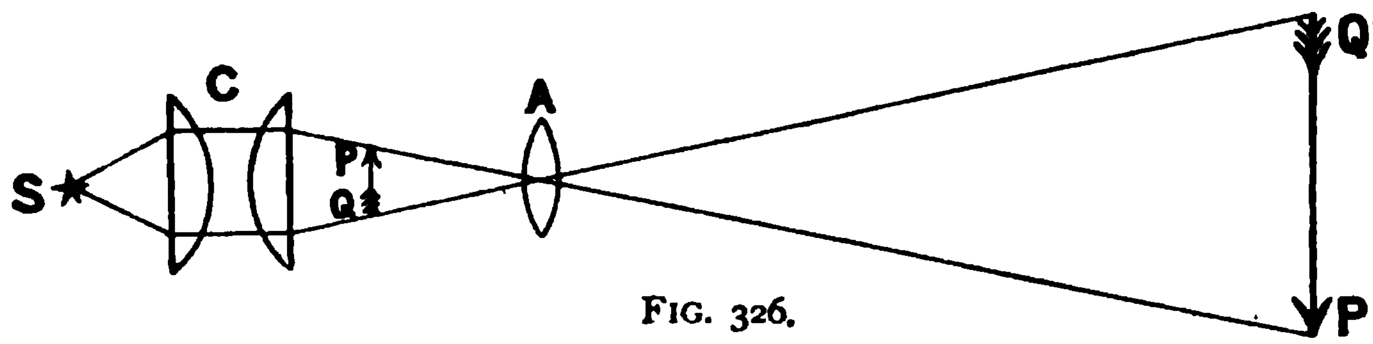


FIG. 326.

receives from the object is spread over a large area, and hence, unless the object is very brightly illuminated, the image will be very feeble.

In front of a brilliant source of light S, such as an electric arc or an oxyhydrogen lime-light, is placed a pair of plano-convex lenses C (Fig. 326), which converge the rays from S, so that they illuminate the object PQ uniformly. A convex lens A is then used to form an enlarged image, P′Q′, of the illuminated object on the screen.

357. Methods of Measuring the Refractive Index.—The most commonly employed method of determining the refractive index of a substance is to take a prism of the substance, if a solid, or a hollow glass prism filled with the substance if it is a fluid, and to measure the angle of the prism θ and the angle of minimum deviation δ, the refractive index being then obtained from the formula

$$\mu = \frac{\sin \frac{1}{2}(\delta + \theta)}{\sin \frac{1}{2}\theta}.$$

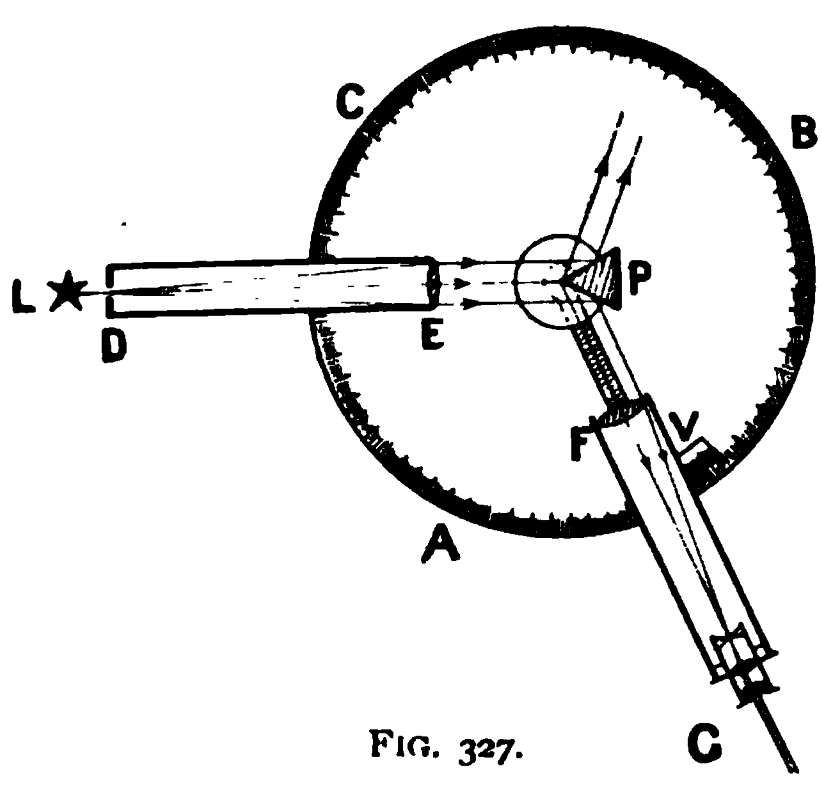


FIG. 327.

The quantities δ and θ can both be measured by means of an instrument called a spectrometer. This instrument consists of a graduated circle ABC (Fig. 327), having a small astronomical telescope FG attached to an arm which can rotate round the centre of the

circle, the position of the telescope being read by means of a vernier V. A tube DE, called the collimator, is fixed radially to the circle, and has a narrow vertical slit D at one end and a convex lens at the other. The distance between the slit and the lens is equal to the focal length of the lens, so that when the slit is illuminated by a source of light L, the rays of light that leave the lens E are all parallel.

The prism P is placed on a small table attached to the circle, and when the refracting angle is being measured, the refracting edge is turned towards the collimator as shown in the figure. The telescope is then turned till the image of the slit, seen by reflection from one face of the prism, coincides with the intersection of two fine cross-wires placed in the eye-piece, and the vernier is read. The telescope is then turned till the image formed by reflection from the other face coincides with the intersection of the cross wires, and the vernier again read. The difference between the vernier readings gives, as shown in § 336, 2θ.

In order to determine δ the prism is removed, and the vernier reading obtained when the telescope is turned so as to see the slit direct, thus obtaining the reading for the direction of the incident rays. The prism is then placed so that the light falls on one of the faces, and the telescope turned so as to catch the deviated rays. By turning the prism round a vertical axis a position is found such that if it is turned in either direction, in order to catch the deviated light, the telescope has to be rotated towards the collimator, or, in other words, the prism is set at minimum deviation by trial. When this has been done, the vernier is read when the deviated image coincides with the cross wires, and the difference between this reading and that for the direct light gives δ.

Another method of determining the refractive index of a medium depends on the fact, as shown in § 343, that a point P at a depth d below the surface of a medium of refractive index, μ, appears, when viewed normally, as if it were at a depth d/μ. In order to apply this method, a microscope C (Fig. 328) is arranged so that it can move in a vertical direction, and the distance through which it is moved can be read on a scale.

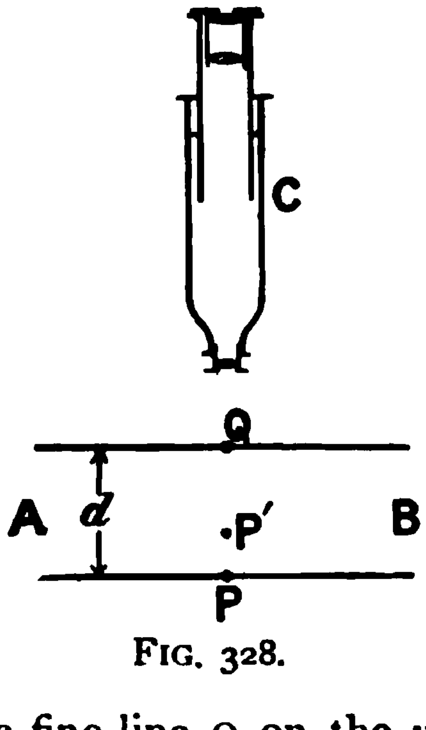


FIG. 328.

The microscope is first focussed on a fine scratch P on the surface of the stage, a plate AB of the material is then introduced, and the microscope again focussed on the scratch P, which will now appear to be at P′, so that the distance through which the microscope will have to be raised will be equal to $\overline{PP'}$. Next, the microscope is focussed on a fine line Q on the upper surface of the slab. The difference between the first and last readings on the vertical scale of the microscope gives d

the thickness of the slab, and the difference between the second and third gives P′Q. Then

$$\mu = \frac{d}{P'Q}.$$

A very convenient method of measuring the refractive index of a substance, particularly for liquids, or when only a small quantity can be procured, depends on the measurement of the critical angle (§ 344) at which total reflection begins. Suppose we have two plates of glass AB and CD (Fig. 329) fixed together in such a way that they include a film of air EF, and that the whole is immersed in a fluid. Let us call the fluid medium 1, the glass medium 2, and the air medium 3. Then if a ray of light travelling in medium 1 is incident at an angle α on the glass, the refracted ray OO′ will be inclined at angle β to the normal. Since AB is a parallel-sided plate, the ray will be incident on the face separating the glass from the air at an angle β.

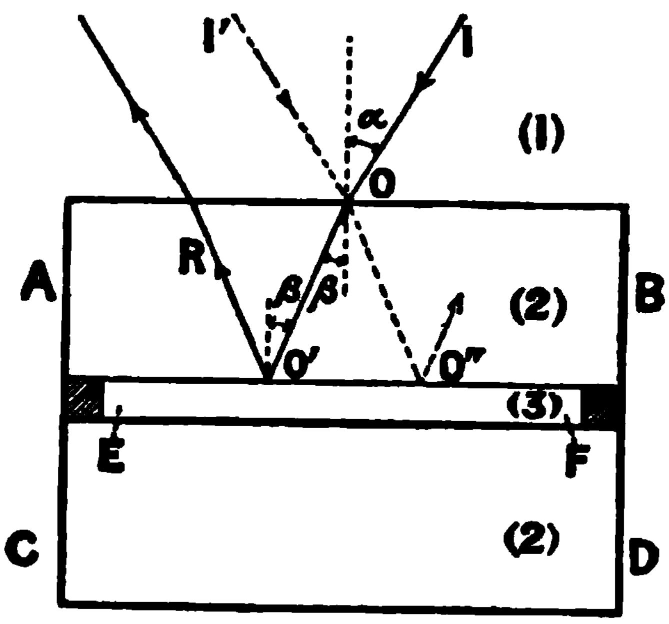


FIG. 329.

If now AB and CD are turned round, *i.e.* the angle of incidence α is altered till the ray OO′ is just totally reflected at O, that is, till β is the critical angle from glass to air, we have the following relations :—

$$\frac{\sin \alpha}{\sin \beta} = {}_1\mu_2,$$

$$\sin \beta = \frac{1}{{}_3\mu_2} = {}_2\mu_3.$$

(See § 344.) Now (§ 342),

$${}_1\mu_2 \times {}_2\mu_3 \times {}_3\mu_1 = 1.$$

Hence substituting $$\frac{\sin \alpha}{\sin \beta} \cdot \sin \beta \cdot {}_3\mu_1 = 1,$$

or $$\sin \alpha \cdot {}_3\mu_1 = 1.$$

Hence $${}_3\mu_1 = \frac{1}{\sin \alpha}.$$

But ${}_3\mu_1$ is the refractive index from medium 3 to medium 1, *i.e.* from air to the liquid, which is what we have called the refractive index of the liquid. Hence if we can measure α we can calculate μ.

To determine the angle α, a parallel beam of light is allowed to traverse a glass-sided trough containing the liquid, and the glass plates, which are attached to a divided circle, are rotated till the light is no longer transmitted through the air film. The position of the plates is then read on the circle, and they are turned in the opposite direction, till the light is again totally reflected, *i.e.* till the light is incident along I′O. The difference between the new reading and the previous one gives twice the angle α.

CHAPTER IV

PHOTOMETRY

358. Illuminating Power and Intensity of Illumination.—It is a matter of common observation that the amount of light emitted by different sources is very different, and before proceeding to describe the methods employed for comparing the amount of light emitted by two sources, we must first consider how we are to define the amount of light emitted by a source. Suppose a point O to be a source of light which is emitting light in all directions. If a sphere be described round O as centre, all the light emitted by the source will fall upon this sphere. Hence the quantity of light which falls on unit area of the sphere will be proportional to the quantity of light emitted by the source.

Thus if we keep the radius of the sphere the same, and use different sources, the quantities of light which fall on unit area will be proportional to the quantities of light emitted by the sources. If the radius of the sphere be unity, the quantity of light which falls on unit area of its surface is taken as the measure of the *illuminating power* of the source. It is evident that, since the area of a sphere of unit radius is 4π, the illuminating power represents $1/4\pi$ of the total light emitted by the source, if the source emits light uniformly in all directions.

The quantity of light which is received by unit area of a given surface, which is placed normally to the incident light, is called the *intensity of illumination* of the surface. If the surface is uniformly illuminated, the total quantity of light received by the surface is equal to the product of the area into the intensity of illumination. If the illumination is not uniform, the intensity of illumination at a point is the quotient of the quantity of light received by an element of area surrounding the point, which is so small that over it the intensity of the light is uniform, by the element of area (*cf.* § 31).

359. The Law of Inverse Square.—Consider a luminous source O, of which the illuminating power in all directions is E, then, if with O as centre we describe a sphere of unit radius, the quantity of light which falls on unit area of this sphere, or the intensity of illumination on the sphere, will be E, and the total quantity of light which falls on the sphere will be $4\pi E$. If now we take round O as centre another sphere of radius r, the first sphere being removed, the whole of the light emitted by the source must fall on the surface of this sphere. Hence, since the total quantity of light emitted by the source is $4\pi E$, and the surface of the

sphere is $4\pi r^2$, and is everywhere at right angles to the incident light, the intensity of illumination on the surface of the sphere is

$$\frac{4\pi E}{4\pi r^2}=\frac{E}{r^2}.$$

Thus the intensity of the illumination varies inversely as the square of the distance from the luminous source. This law is known as the inverse square law.

Since light is propagated in straight lines, the inverse square law holds if the illuminating power of the source is different in different directions.

Thus suppose O (Fig. 330) is a luminous source, and ABCD is a small square aperture, each side of which is of length l, in a screen placed at a distance d_1 from the source, and the light which passes through this aperture is received on a screen A′B′C′D′ at a distance d_2 from the source. If ABCD and A′B′C′D′ are at a considerable distance from the source, compared to their area, the intensity of illumination will be practically uniform on each, and the light which would fall on ABCD, if the aperture were filled up, now falls on A′B′C′D′. Since light is propagated in straight lines, we have, from the similar triangles BOC and B′OC′,

FIG. 330.

$$\frac{\overline{B'C'}}{\overline{C'O}}=\frac{\overline{BC}}{\overline{CO}},$$

or

$$\overline{B'C'}=\frac{ld_2}{d_1}.$$

Hence the area A′B′C′D′ is

$$\frac{l^2d_2^2}{d_1^2}.$$

But the area of ABCD is l^2. Hence if I is the quantity of light which falls on ABCD or A′B′C′D′, we have

$$\frac{\textit{Intensity of illumination of } ABCD}{\textit{Intensity of illumination of } A'B'C'D'}=\frac{I/l^2}{I\Big/\frac{l^2d_2^2}{d_1^2}}$$

$$=\frac{d_2^2}{d_1^2}.$$

So that, as before, the intensity of illumination is inversely as the square of the distance; and in this case we may take the area ABCD as small as we please, so that the law applies to the light emitted from a source in

any given direction, whether the illuminating power is the same in all directions or not.

360. Unit of Illuminating Power.—We have hitherto spoken of the quantity of light emitted by a source or received by a screen, but have not said anything as to what measure of this quantity is used. As we have no means of measuring this quantity absolutely, we have to adopt some standard source of light, and say that its illuminating power is unity. The standard ordinarily in use in this country, although a most unsatisfactory and variable one, is the illuminating power of a "standard candle." The standard candle is a sperm candle, of which six weigh a pound, which burns 120 grains (7.776 grams) of wax in the hour. The candle power of any source is then the number of standard candles which would have the same illuminating power as the given source. Other standards of illuminating power are the Carcel and the Hefner-Alteneck lamp. The Carcel is a lamp burning colza oil at a fixed rate, and the Hefner-Alteneck a lamp burning amyl-acetate, the height of the flame being adjusted to a fixed value. Of these the latter is much more constant than either the Carcel or the standard candle. Violle has proposed to use, as a standard, the light emitted normally by a square centimetre of platinum at the temperature of its melting-point, but this can hardly be described as a practicable standard for general use.

361. Photometry.—Although the eye is only capable of very roughly estimating the relative intensity of the illumination of two surfaces, yet it is capable of telling with considerable accuracy when two adjacent surfaces are illuminated with equal intensity. Hence, to compare the illuminating powers of two sources, they are so arranged that two adjacent patches on a screen, each patch being illuminated by only one of the sources, appear illuminated with the same intensity. Now if I_1 and I_2 are the illuminating powers of the sources, and d_1, d_2 the distances between each source and the patch of the screen which *it illuminates* when the intensity of illumination of the two patches is the same, we have that the intensity of the illumination due to one source at its patch is I_1/d_1^2, and that due to the other at its patch is I_2/d_2^2. Hence, since these intensities are equal, we have

$$\frac{I_1}{d_1^2}=\frac{I_2}{d_2^2},$$

or

$$\frac{I_1}{I_2}=\frac{d_1^2}{d_2^2},$$

or the illuminating powers of the two sources are directly as the square of the distances at which they respectively produce equal intensities of illumination.

There are a number of arrangements in use for obtaining the two patches, illuminated each by a separate source, in such relative positions that any difference in their intensities of illumination may most easily be

detected by the eye. One of the simplest *photometers*, as the instruments used for comparing the illuminating powers of two lights are called, is that due to Rumford.

A Rumford's photometer consists of an upright screen AB (Fig. 331), which is covered with white unglazed paper—white blotting-paper does very well—and in front of which an upright opaque rod R about an inch in diameter is placed. There are two scales, M and N, inclined at the same angle to the screen, along which the two sources, P and Q, can be moved. If the source P only were present, the rod R would cast a shadow *fd* on the screen, while if the light Q only were present, the shadow of the rod would be at *af*. Hence when both lights are present, while the parts A*a* and *d*B of the screen are illuminated by the two sources, the part *fd*, which receives no light from P owing to the interposition of the opaque rod, is only illuminated by Q, and the part *af* is only illuminated by P. If, then, the distances of the lights from the screen are adjusted till the two patches *fd* and *af* are equally illuminated, we shall have that the

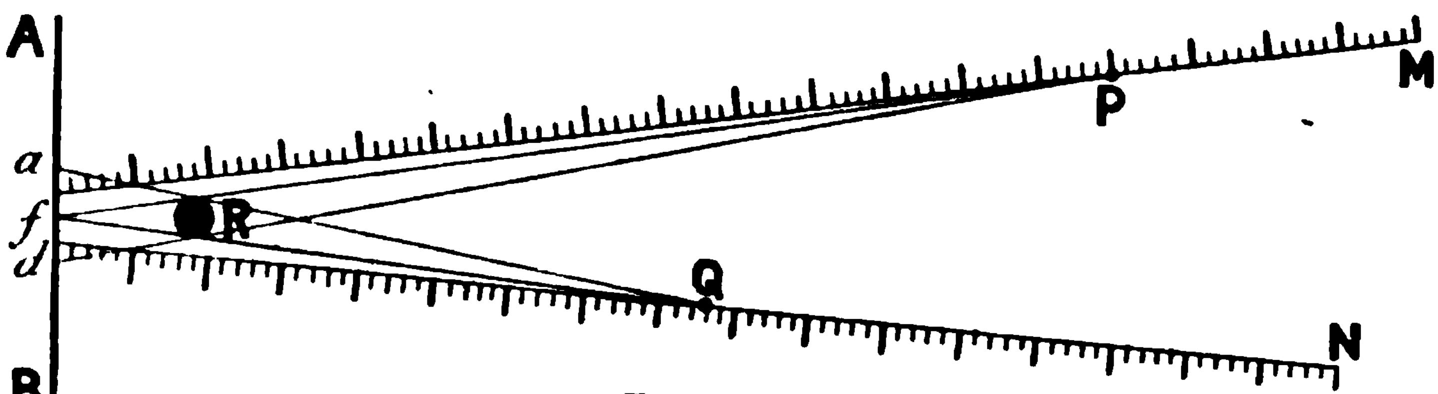


FIG. 331.

source P produces, at the distance of P from *af*, the same intensity of illumination as the source Q produces at the distance of Q from *fd*. Hence if d_1 is the distance of P from the part of the *screen* which it only illuminates, namely *af*, and d_2 is the distance of Q from the part of the screen it only illuminates, namely *fd*, we get, if I_1 and I_2 are the illuminating powers of P and Q respectively, that

$$\frac{I_1}{I_2} = \frac{d_1^2}{d_2^2}.$$

The distance of the rod R from the screen is adjusted so that the shadows of the rod cast by the two lights just touch one another, as it is found that the eye can best judge when they are equally illuminated under these circumstances.

Another form of photometer which is frequently used is Bunsen's grease-spot photometer. This photometer consists of a small screen which has a central spot of grease, or in some other way is constructed so that the central portion is more translucent than the surrounding parts. If such a screen is held between the eye and a source of light, more of the light passes through the grease-spot than through the sur-

rounding more opaque parts of the screen, so that the spot appears brighter than the surrounding paper. If, however, the screen is held against a dark background and illuminated from the front, the grease-spot will appear dark, for more of the light which is incident on the screen is transmitted through the spot than through the rest of the screen, and hence less is reflected or diffused so as to reach the eye by the spot than by the surrounding parts. If the screen is equally illuminated on both sides, then the spot diffuses less of the light received from the one source than the surrounding parts, but it transmits more of the light from the other source, so that these two effects just neutralise one another, and the spot appears of the same brightness as the surrounding paper.

The screen with the grease-spot is placed between the two sources whose intensities have to be compared, and moved about till the grease-spot can no longer be distinguished from the rest of the screen. If, when this adjustment has been made, the distances of the two sources from the screen are d_1 and d_2, we have, as before,

$$\frac{I_1}{I_2}=\frac{d_1^2}{d_2^2}$$

In using Bunsen's photometer, it is of assistance if both sides of the screen can be seen simultaneously. The usual arrangement employed to secure this end is a system of two mirrors inclined at 45° to the screen. With this arrangement one side of the screen is seen with one eye and the other with the other, so that if, as is generally found to be the case, one eye is more sensitive than the other, a wrong setting may be made. This source of error is removed in the Lummer-Brodhum photometer, which consists of an opaque screen AB (Fig. 332), each side being illuminated by one of the sources which are to be compared. The two sides of the screen are viewed by means of two plane mirrors, M_1 and M_2, and a double glass prism, CD. This prism consists of two right-angled prisms, the longest face of one being partly bevelled away, fastened together with Canada balsam Owing to total internal reflection (§ 344), the central rays reaching an eye at E come from the left-hand face of the screen, and the surrounding rays come from the right-hand side, as is shown in the figure.

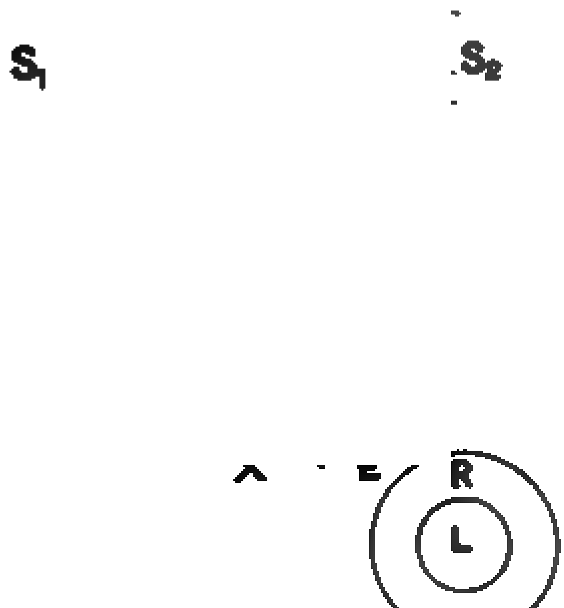


Fig. 332.

Hence the observer moves the photometer till the central patch L and circumferential parts R appear of the same brightness, when the intensity of the illumination on the two sides of the screen is the same; the relative powers of the two sources is then as the square of their distances from the screen.

The relative illuminating powers of two sources is only strictly comparable when the colour of the light emitted is the same for both. If the colours differ, we can only compare their illuminating power for the lights of different wave-length which are included in the light given by *both* sources. In order to perform this comparison, a spectrum (§ 367) is formed with the light from each source, and the intensities of the different portions of the spectra which are common to both are compared. A rough comparison can be made by comparing the powers of the two sources, when a red, a yellow, and a blue-coloured glass is placed in turn between each source and the photometer screen. The three values for the relative illuminating powers thus obtained will give an idea as to the relative composition of the light given by the two sources.

CHAPTER V

VELOCITY OF LIGHT

362. Finite Velocity of Light—Römer.—An entirely new era in the history of the science of light was introduced by the Danish astronomer Römer in 1676, when he not only showed that light did not travel instantaneously, as had been previously supposed, but also measured the velocity with which light travels through interplanetary space.

The planet Jupiter has four moons, and as these revolve round the planet they disappear once in each revolution, for when they pass into the shadow of the planet cast by the sun they become invisible, and are said to be eclipsed, for we only see the planets and their satellites by the light of the sun which they reflect.

If any one of Jupiter's moons revolves round the planet with a uniform angular velocity, as is the case with our moon, then the time which elapses between one passage of the moon into the shadow and the next ought to be constant, since it would be equal to the periodic time of the moon's revolution round the planet. However, if the times between successive eclipses of the planet are noted, as seen from the earth, it is found that they are not all equal.

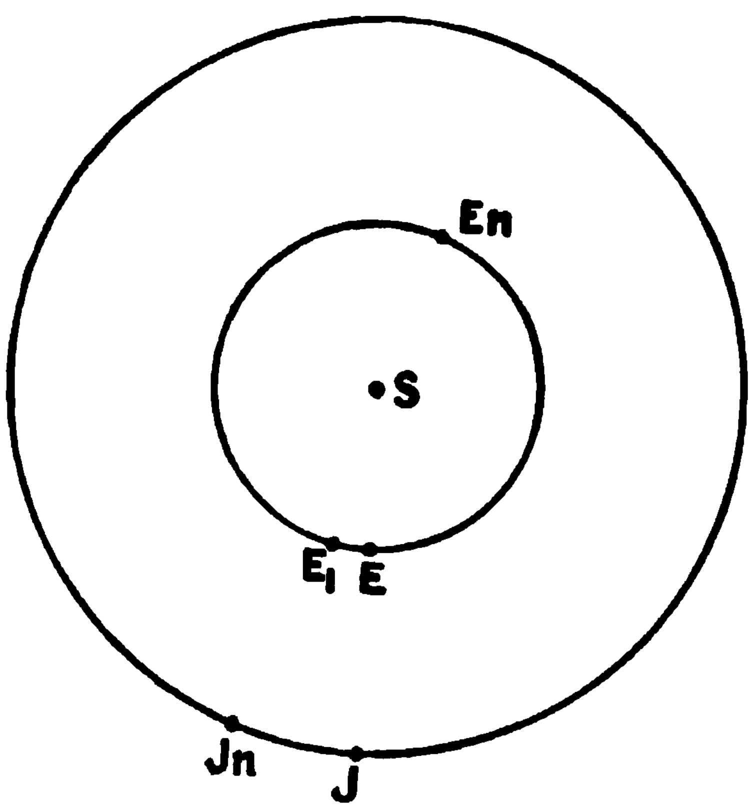


FIG. 333.

Römer accounted for this phenomenon by supposing that light took an appreciable time to travel from Jupiter to the earth. Let S (Fig. 333) be the sun, E the earth, and J Jupiter; and suppose that when an eclipse of one of the moons occurs these three are in the relative positions shown, the earth being at its nearest point to Jupiter. When the next eclipse occurs, the earth will have moved round in its orbit to some such position as E_1, while Jupiter will have only moved a short distance round in its orbit. Hence the distance between the earth and Jupiter is now greater than before. If V is the speed

with which light travels, then the observed time at which the eclipse occurs when the earth is at E will be at a time $\overline{EJ}/V$, *i.e.* the time taken by the light to traverse the space $\overline{EJ}$, after the actual eclipse. When the earth is at E_1, the observed time will be $\overline{E_1J}/V$ later than the actual time. Hence, if θ is the actual time between two successive eclipses, the observed interval is $\theta+\overline{E_1J}/V-\overline{EJ}/V$, and since $\overline{E_1J}$ is greater than $\overline{EJ}$, the observed interval will be greater than the true interval; since, however, we cannot observe the *true* interval θ, the quantity V cannot be calculated from the observation of a single interval in this way.

At each successive eclipse the earth will be further and further away from Jupiter, till finally they come into the positions J_n, E_n, when they are at their maximum distance apart. After this, at each successive eclipse the distance will diminish, and hence the observed interval be less than the true interval.

Suppose that n eclipses occur between that which occurs when Jupiter and the earth are nearest together (at conjunction), and that which occurs when they are at their greatest distance (opposition), the actual interval between the first and last of these eclipses is $n\theta$. The observed interval is $n\theta+\overline{E_nJ_n}/V-\overline{EJ}/V$, or if d is the diameter of the earth's orbit, so that $\overline{E_nJ_n}-\overline{EJ}=d$, the observed interval T_1 is $n\theta+d/V$.

The actual interval between the eclipse when the earth is at opposition and the one when it is again at conjunction will also be $n\theta$. The observed interval, T_2, since at the end of the series the earth is nearer Jupiter than at the commencement by a distance d, will be $n\theta-d/V$.

Thus
$$T_1=n\theta+\frac{d}{V},$$
and
$$T_2=n\theta-\frac{d}{V};$$
$$\therefore \quad T_1-T_2=\frac{2d}{V},$$
or
$$V=\frac{2d}{T_1-T_2}.$$

Hence if we know the diameter, d, of the earth's orbit, we can calculate the velocity of light, V, from the difference between the interval T_1 which elapses between an eclipse at conjunction and the eclipse at the next opposition, and the interval T_2 between this eclipse at opposition and the one which occurs at the next conjunction.

The moon chosen was the innermost, which makes a revolution in about $1\frac{3}{4}$ days, and it was found that, starting at conjunction, the interval between the first eclipse of this planet and the 113th (when the earth and Jupiter came into opposition) exceeded the interval between the 113th and the 225th (when the earth and Jupiter were again in conjunction) by 33.2 minutes, and hence $T_1-T_2=33.2$ minutes. If d, or the diameter of

the earth's orbit, is taken as 195,600,000 miles, or 298,600,000 kilometres, this gives 186,300 miles per second, or 299,800 kilometres per second as the velocity of light.

363. Fizeau's Method of Measuring the Velocity of Light.—The accuracy of the determination of the velocity of light by Römer's method is limited by the accuracy with which we know the diameter of the earth's orbit, hence it is important to determine the velocity of light between two terrestrial points, the distance between which can be directly measured. The first to perform this experiment was Fizeau, who in 1849 measured the velocity of light by using a method depending on the eclipsing of a source of light by the teeth of a rapidly rotating wheel, the principle of the experiment resembling Römer's method.

A bright source of light was placed at L (Fig. 334), and after passing through a lens A, a certain proportion of the rays of light was reflected from the surface of an unsilvered plate of glass, G, placed at an angle of 45°. The reflected rays came to a focus at F, this point being the

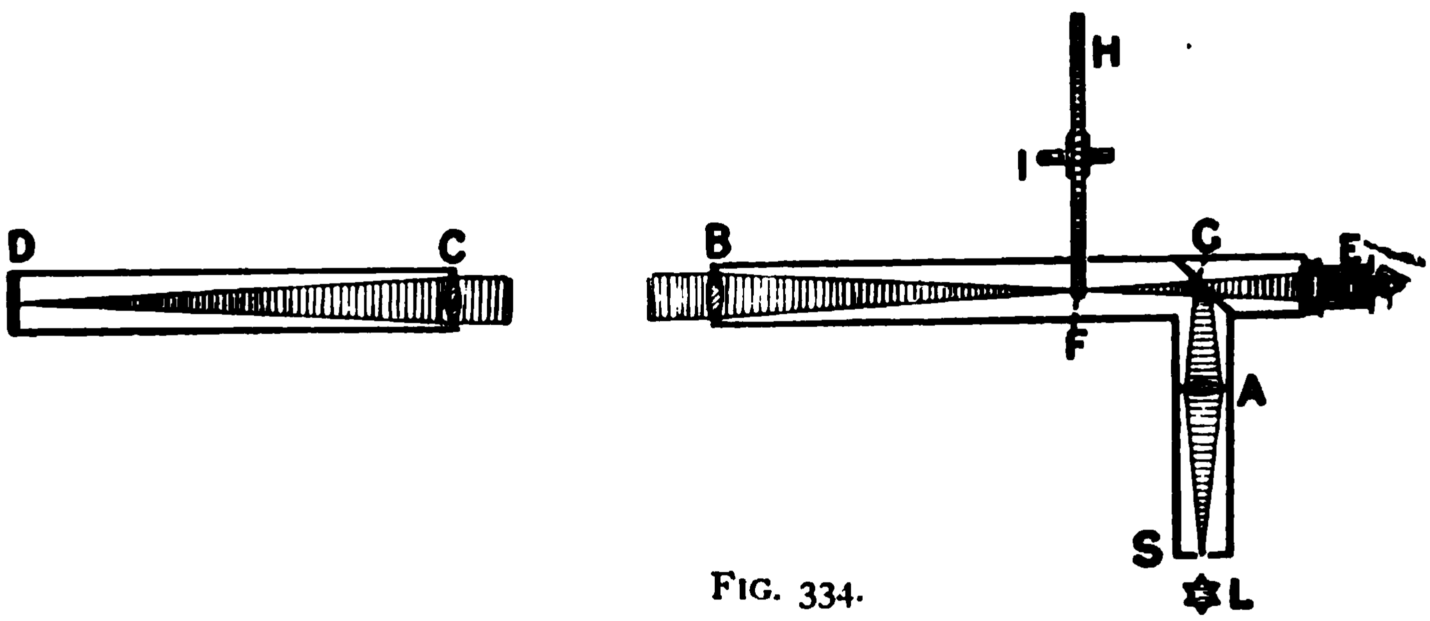


FIG. 334.

principal focus of a second lens B. Thus the light left B in a parallel beam, which, after traversing a distance of about four miles, fell on a lens C, and was brought to a focus at the surface of a spherical mirror D. The curvature of this mirror is such that the lens C is at its centre of curvature, and hence the rays are reflected back along their path, so that on emerging from the lens C they again form a parallel beam. This reflected beam falls on the lens B, is brought to a focus at F, and then falls on the plate of glass G, where some of the rays will be reflected, and some will be transmitted and enter the eye-piece E, so that a bright star will be seen by the observer, formed by light which has travelled to D and back again. A toothed wheel, H, which can be rapidly rotated round an axle, I, is so arranged that when a tooth passes F the light is intercepted, but when a space passes F the light is allowed to pass.

If the wheel is slowly rotated, an observer at E will see a bright star when a space passes F, while when a tooth passes there will be darkness, so that as the wheel rotates the star alternately appears and dis-

appears; but if the speed is such that more than twenty teeth pass per second, owing to the persistence of vision, a permanent star will be seen.

If light took no time to travel from F to D and back again, then all the light that passed through any space would be able to pass back again through the same space, since by supposition the light takes no time to travel from F to D and back, and hence the wheel would not have moved between the starting of the light and its arrival back at F. If, however, the light takes an appreciable time to travel, then, as the speed of the wheel is gradually increased, it will eventually rotate so fast that by the time the return light reaches F the wheel will have turned so that a tooth will have moved round, and will occupy part of the space which was occupied by a space when the light started, so that part of the return light will be cut off by the tooth, and hence the star seen at E will be of decreased brightness. As the speed is further increased, more and more of the return light will be intercepted by the succeeding tooth, till finally all the light which gets through a space is, on its return, intercepted by the succeeding tooth, and no star is seen. If the speed is yet further increased, the returning light will begin to get through the space next after the one through which it passed on its way out, and a bright star will again be seen. Hence, as the speed is increased, the star alternately appears and vanishes.

If the wheel contains d teeth, then the angle ACB (Fig. 335) subtended by the interval between two consecutive teeth or spaces at the centre will be $360°/d$ or $2\pi/d$ in circular measure. Hence half this angle, or π/d, is the angle through which the wheel has to turn so that a tooth may exactly occupy the position previously occupied by a space.

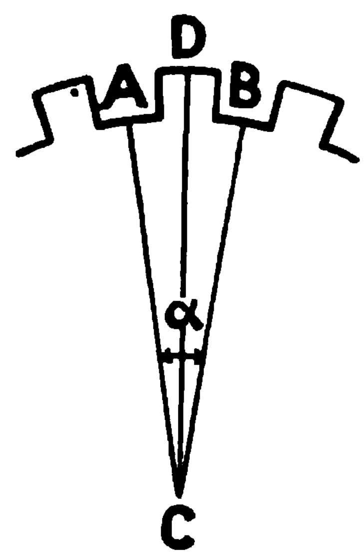


FIG. 335.

If the wheel makes n revolutions per second when the first eclipse occurs, the angle swept out by any radius AC in one second will be $2\pi n$. Hence the time taken to turn through the angle π/d will be $\frac{\pi}{d}\Big/2\pi n$, or $1/2dn$. If l is the distance between F and D, the distance passed over by the light while the wheel has turned through an angle π/d is $2l$. Hence the light has travelled a distance $2l$ in a time $1/2dn$, or the velocity of light V is given by

$$V = \frac{2l}{\frac{1}{2dn}} = 4ldn.$$

In one of the experiments, l was equal to 8633 metres, the wheel had 720 teeth, and when the star was first eclipsed it made 12.6 revolutions per second, so that

$$V = 4 \times 8633 \times 720 \times 12.6 = 313274304 \text{ metres per second.}$$

More recent experiments made by Cornu, using this method, gave 300,400 kilometres per second, or 186,662 miles per second, as the velocity of light.

364. Foucault's Method of Measuring the Velocity of Light.—In the year 1850, Foucault succeeded in measuring the time light took to travel over a distance of about twenty metres. His method consists in causing a beam of sunlight to fall on a slit S (Fig. 336), by means of a heliostat. The light transmitted by S passes through an unsilvered glass plate G, falls on a convex lens A, and then on a plane mirror B, which can be rapidly rotated round an axis perpendicular to the plane of the figure. For one position of the mirror B, the reflected light falls upon a second mirror C. This latter is a concave mirror, the radius of curvature being equal to $\overline{BC}$. Hence, if the mirror B is at rest, the

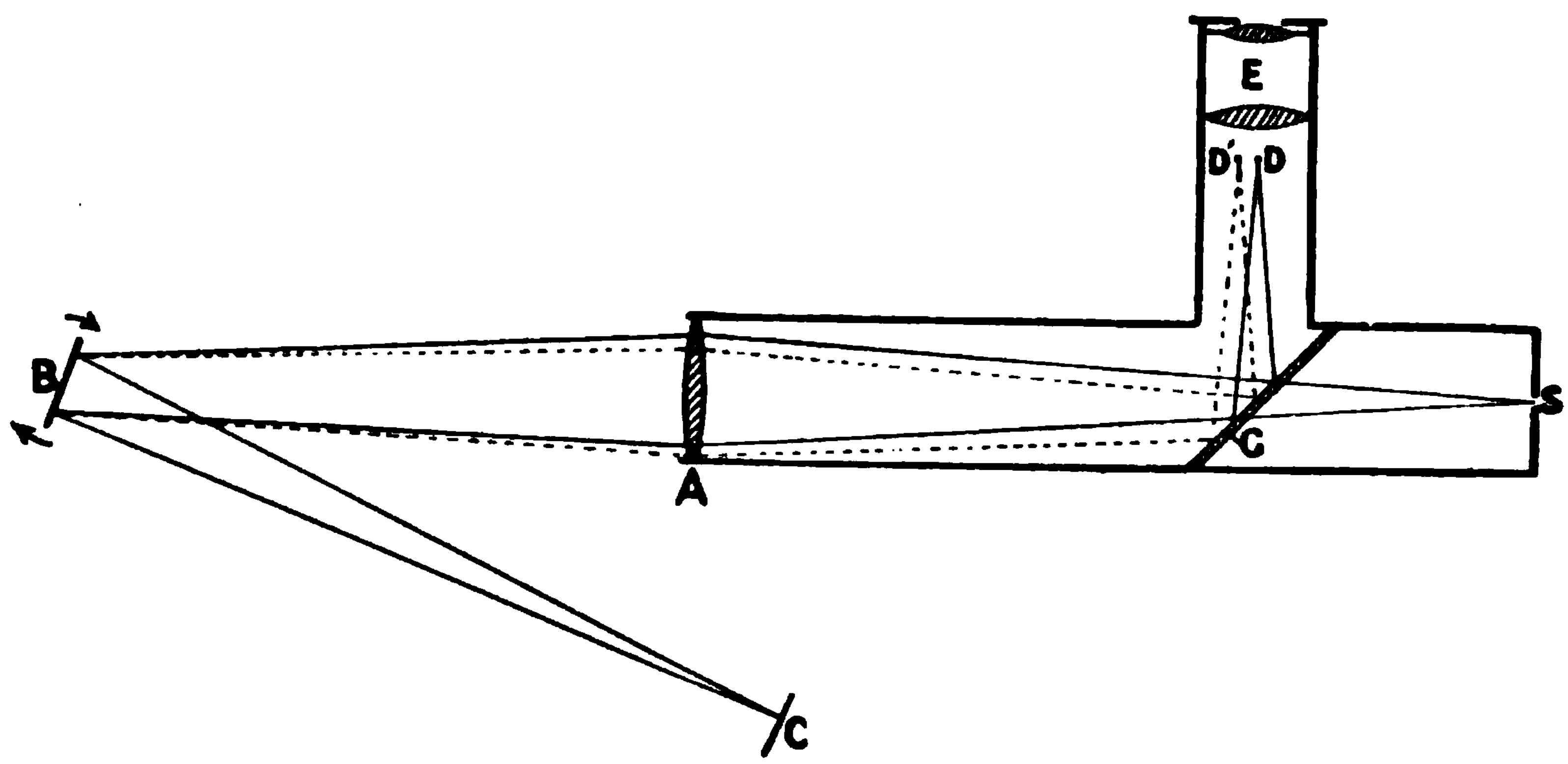


FIG. 336.

light reflected at C will retrace its path, being partly reflected at G, so as to form an image of the slit S at D, this image being observed through an eye-piece E. If the mirror B is rotated, the light is only reflected back from C once in each revolution, if the mirror is only silvered on one side. If, further, during the time taken for the light to travel from B to C and back to B, the mirror has appreciably turned, the return rays will be reflected by B in a slightly different direction to that which they would have taken had the mirror been at rest, and the image D will be displaced, as shown by the dotted lines, to D′, the amount of the displacement being read off on a scale placed at D. In order to count the speed of rotation of the mirror B, which was driven by a small steam-turbine, Foucault placed a toothed wheel so that the teeth were illuminated by the intermittent beam of light reflected from the rotating mirror. If the wheel was rotating at such a speed that during

the interval between two flashes one tooth had just moved into the position occupied at the previous flash by the preceding tooth, then the wheel would appear to be at rest. When this is the case, the time taken by the mirror to make one revolution is equal to the time taken by the wheel to turn through the angle ACB (Fig. 335), that is, $1/dn$, where d is the number of teeth in the wheel, and n the number of revolutions it makes in a second. Hence the number N of revolutions made by the mirror in one second is dn.

Suppose that the mirror is at rest, the image being at D, and that, when the mirror is turned through a small angle α, the image is moved to D' through a distance a. If the distance between the mirror B and the image D (*i.e.* the distance $\overline{BG}+\overline{GD}$) is called r, then the movement of the image through a distance a corresponds to a rotation of the reflected rays forming this image through an angle (measured in circular measure) of a/r. Further, since, as has been shown in § 331, the angle through which the reflected rays are turned is twice the angle through which the mirror has been turned, we have

$$2\alpha=\frac{a}{r},$$

or

$$\alpha=\frac{a}{2r}.$$

If the displacement of the image when the mirror is rotating rapidly is a, this means that during the time the light has taken to travel from B to C and back, the mirror has turned through the angle α. Since the mirror makes N revolution per second, it will turn through an angle α in a time $\alpha/2\pi N$; or substitute for α its value, obtained in terms of a and r, in a time $a/4\pi Nr$. During this time the light has travelled a distance $2l$, if l is the distance between B and C, so that the velocity of light is given by

$$V=\frac{2l}{\dfrac{a}{4\pi Nr}}=\frac{8\pi Nrl}{a}.$$

Using this method, Michelson has obtained 299,853 kilometres per second as the velocity of light, with a possible error of ± 60 kilometres.

365. Aberration.—A calculation of the velocity of light was made in 1727 by Bradley from the apparent changes which take place in the observed positions of the fixed stars. The simplest method of explaining the principle of this method of calculating the velocity of light is to consider an analogous case. Suppose that a shot is fired from a cannon C (Fig. 337) against a ship AB, which is moving rapidly at right angles to the direction of the trajectory of the shot. If the shot enters the ship at the point D, it will not leave the ship at the point E, for while the shot is travelling across the ship this latter will have moved forward, so

that some such point as F will now be in the line CD of the trajectory, and hence the shot will leave the ship at F. Now, as far as the track of the shot left in the ship is concerned, it appears to have come from the point C′, or the cannon is apparently displaced in the direction in which the ship is travelling.

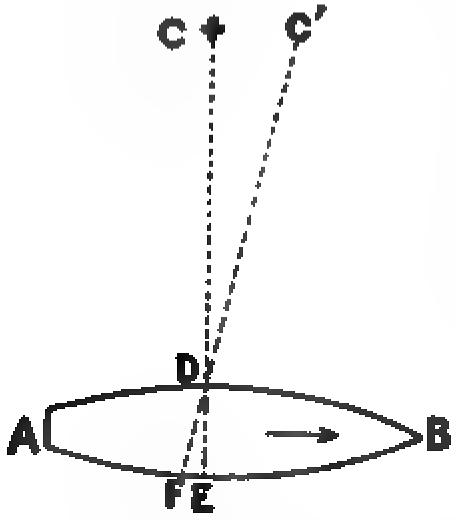


Fig. 337.

Now consider the ship to be replaced by the earth moving in its orbit round the sun, and the shot to be replaced by a light-wave reaching the earth from a star, the line joining the earth and star being at right angles to the direction of motion of the earth. If the earth were at rest, and a telescope D (Fig. 338) were pointed at the star, the axis of the telescope prolonged would pass through the star C. When the earth is moving, however, the telescope is appreciably moved forward in the earth's orbit in the time between the light-waves reaching the objective D and the eye-piece; and hence, if the axis of the telescope is still directed so as to pass through the star, the waves of light will be left behind and form an image of the star at F to one side of the cross wires. Just as in the case of the boat, the shot, which if the boat were at rest would have reached the point E, hits the point F. Thus in order to bring the image of the star on to the cross wires, the axis of the telescope must be pointed along FC′. Now in the time the light has taken to travel from D to F the earth has moved through the distance FE. Hence if the angle CDC′, which is called the aberration constant, is α, and V and v are the velocity of light and of the earth respectively, we have—

$$v/V = \tan \alpha,$$
$$\text{or } V = v/\tan \alpha.$$

Hence if v and α are known, V, that is, the velocity of light, can be calculated.

Fig. 338.

The above explanation of the phenomenon of aberration is that commonly accepted; it is, however, by no means satisfactory. Thus the quantity V is the velocity of light in the medium, filling the telescope tube, so that if the tube is filled with water, in which the velocity of light is less than in air, we should expect the aberration constant α to increase. No such effect is, however, observed. The discussion of the various theories which have been propounded to explain this discrepancy would lead us beyond the scope of this book, and so

we must content ourselves with having drawn attention to the existence of the difficulty.

366. Theories as to the Nature of Light.—We have hitherto assumed that light consists of a wave-motion of some kind, and we now have to consider the evidence on which this assumption is founded.

There have been two principal theories of light. In one of these, which was adopted by Newton, and is called the *emission theory*, a luminous body is supposed to be emitting small particles, called light-corpuscles, which travel out in all directions in straight paths, and all with the same velocity. These light-corpuscles were supposed to cause the sensation of light by their impact on the retina. Since light can traverse not only empty space, but also some forms of matter, these corpuscles must be able to travel through space, and also through matter, which they were supposed to do by passing between the molecules. On this theory the rectilinear propagation of light and the formation of shadows at once followed.

In order to explain the law of refraction on the emission theory, it was assumed that when a corpuscle came near the surface of separation between two media, it was attracted by the denser medium.

Thus let AB (Fig. 339) represent the line of demarcation between two media, say air and water, and $\overline{IO}$ the path of an incident corpuscle. If v_1 is the velocity with which the corpuscle travels in air, then we may resolve this velocity into two components, one, $\overrightarrow{IN}$, parallel to the surface of separation of the media, and the other, $\overrightarrow{IM}$, perpendicular to this surface. If the water particles exert an attraction on the light-corpuscle when it gets near the surface of separation, this force must act normally to the surface, *i.e.* along ON′. Hence, while the vertical component of the velocity of the corpuscle is increased, the horizontal component, $\overrightarrow{IN}$, will remain unchanged. If then we take $\overline{OM'}$ equal to $\overline{OM}$, and ON′ equal to the increased vertical component, and complete the parallelogram, $\overrightarrow{OR}$ will represent the velocity of the corpuscle in the water, and its path, that of the refracted ray, will be along OR. From Snell's law we have—

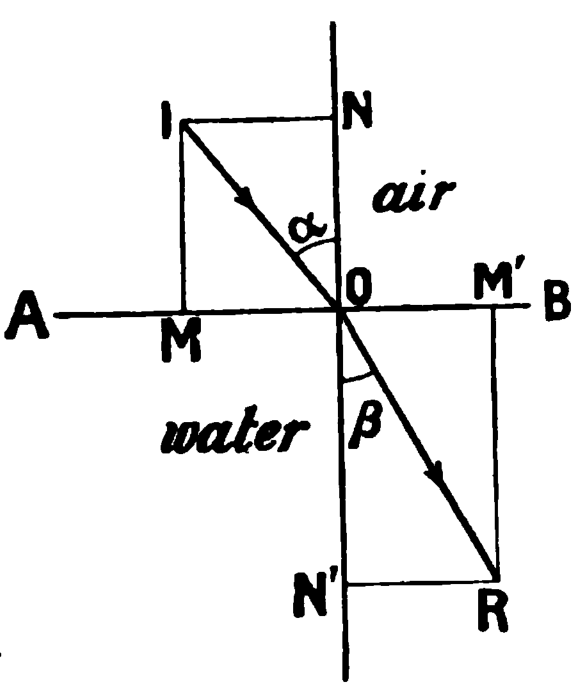


FIG. 339.

$$ {}_a\mu_w = \frac{\sin \alpha}{\sin \beta} = \frac{\overline{IN}}{\overline{IO}} \div \frac{\overline{N'R}}{\overline{RO}} = \frac{\overline{RO}}{\overline{IO}}. $$

Since $\overline{IN} = MO = OM' = N'R.$

But $\overline{IO}$ represents the velocity of the corpuscle in air, while $\overline{RO}$ represents its velocity in water. Hence

$$\frac{\textit{Velocity in water}}{\textit{Velocity in air}} = {}_a\mu_w,$$

and since the refractive index from air to water is greater than unity, it follows that, according to the emission theory, the velocity of light must be greater in water than in air.

In the second theory, called the undulatory or wave theory of light, a luminous body is supposed to set up vibrations in an all-pervading ether, and these vibrations are supposed to travel through the ether, and when they enter the eye excite the sense of vision. During the passage of the light from the source to the eye, the energy emitted by the source, and which we recognise when it is given up to the retina as light, must be stored up in the ether.

On the older undulatory theory, it was supposed that light-waves consisted of a transverse vibratory movement of the ether itself, but a difficulty was introduced by the fact that, if we suppose that the motion is propagated by the successive parts of the ether setting each other in motion by mutually attracting forces, these forces would be inclined to the direction in which the wave was travelling, and hence they would have a component in the direction of the wave normal, and this component would tend to set up longitudinal waves, in addition to the transverse waves which are required to explain optical phenomena. We have no evidence, however, of the existence of such longitudinal waves in the ether. In the later form of the undulatory theory, called the electromagnetic theory of light, the supposition is made that the vibrations consist not in the change in position of the ether particles, but in a periodic alteration in the electrical and magnetic condition of the ether during the passage of the light. This supposition does not lead to the same difficulty as to the formation of longitudinal waves as does the older theory, and hence possesses a marked advantage.

Since both forms of the theory suppose the existence of a transverse vibration set up in the medium, and only differ as to the nature of the entity the displacement of which constitutes the vibration, the explanations which we shall make in the succeeding sections, since they do not involve the *nature* of the waves, will apply to either form of the theory. We shall also sometimes talk of the displacement of an ether *particle* during the vibration, but this must be taken as a short and convenient method of stating the displacement of the electric and magnetic condition of the ether at the point under consideration.

We now pass on to consider what assumptions as to the relative velocity of light in air and in water have to be made on the undulatory theory to account for the refraction of light when it passes from air to

water. Let AB (Fig. 340) be the line of separation between air and water. Let PP′ represent a wave-front in the air, then if v_a is the velocity of light in air, the time taken for the point P′ on the wave-front to reach the second medium will be $\overline{P'O}/v_a$. During this time the point P on the wave-front will have travelled into the water, and if v_w is the velocity in water, it will have travelled a distance $\frac{\overline{P'O}}{v_a} \times v_w$. If then, with centre P and radius equal to this

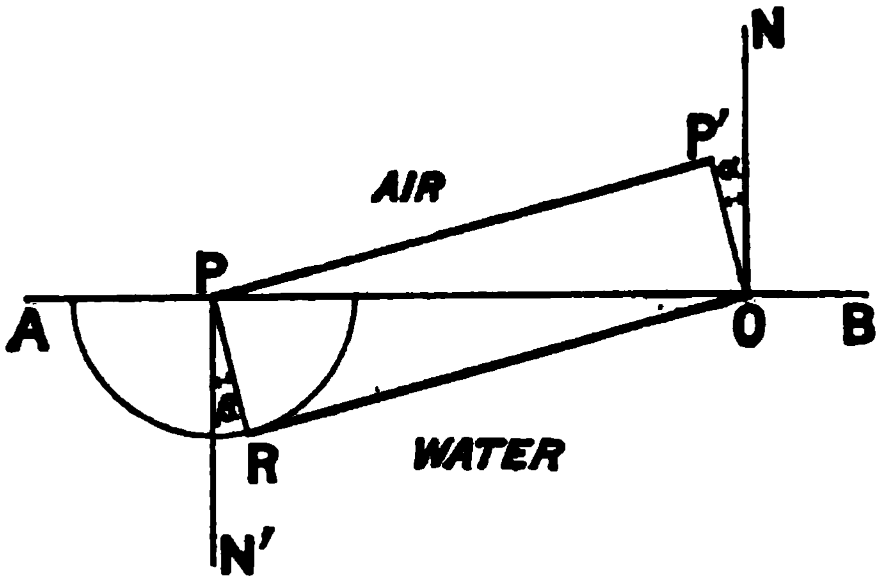


FIG. 340.

distance, we describe a circle, and then from O draw a tangent $\overline{OR}$, $\overline{OR}$ will represent the wave-front in water at the instant when the point P′ on the wave-front PP′ reaches O (see § 273). If ON and PN′ are normals to AB, we have

$$_a\mu_w = \frac{\sin NOP'}{\sin N'PR}.$$

Now in the triangle PP′O the angle at P′ is a right angle, hence the two angles P′PO and P′OP are together equal to a right angle. But the angles P′ON and P′OP are also together equal to a right angle. Hence the angle P′ON is equal to the angle P′PO. In the same way, the angle N′PR is equal to the angle POR. So that

$$_a\mu_w = \frac{\sin NOP'}{\sin N'PR} = \frac{\sin P'PO}{\sin POR} = \frac{\overline{P'O}}{\overline{PO}} \div \frac{\overline{PR}}{\overline{PO}} = \frac{\overline{P'O}}{\overline{PR}}.$$
$$= \overline{P'O} \div \frac{\overline{P'O}}{v_a} v_w = \frac{v_a}{v_w}.$$

Or the velocity of light in air is to that in water in the ratio of the refractive index from air to water. It will thus be seen that according to the undulatory theory light travels slower in water than in air, while according to the emission theory it travels more quickly in water. Thus a measurement of the velocity of light in air and in water would form a crucial experiment to test the validity of the rival theories. This crucial experiment was performed by Foucault, who placed a tube filled with water and closed by glass ends between the fixed mirror C and the rotating mirror B, and thus was able to measure the velocity of light in water, and found it to be *less* than in air. This experiment, although it does not in any way *prove* the truth of the undulatory theory, yet shows that the emission theory, at any rate, cannot be true.

CHAPTER VI

DISPERSION

367. Dispersion.—The phenomenon of refraction is not in reality as simple as we have hitherto considered it to be, for if a narrow parallel pencil of white light, such as sunlight, is allowed to pass obliquely from one medium to another, it is found that in the second medium the white light is split up into light of several colours, a phenomenon which is referred to as *dispersion*.

Thus if a beam of parallel rays of white light, such as is obtained by reflecting sunlight through a narrow slit, is introduced into a dark room and meets a screen DE at F, forming a white patch of light, then on interposing a prism ABC (Fig. 341) in the path of the beam with its refracting edge parallel to the slit, the light will be refracted towards the base of the prism, but the patch on the screen is no longer the same size as before, nor is it white. The patch is drawn out in the direction RV, in which the light is deviated, and exhibits all the colours of the rainbow. These colours pass imperceptibly the one into the next, but starting with red nearest the original undeviated patch F, the colours pass through orange, yellow, green, blue, indigo, to violet, which is the most deviated. These colours constitute what is called a spectrum.

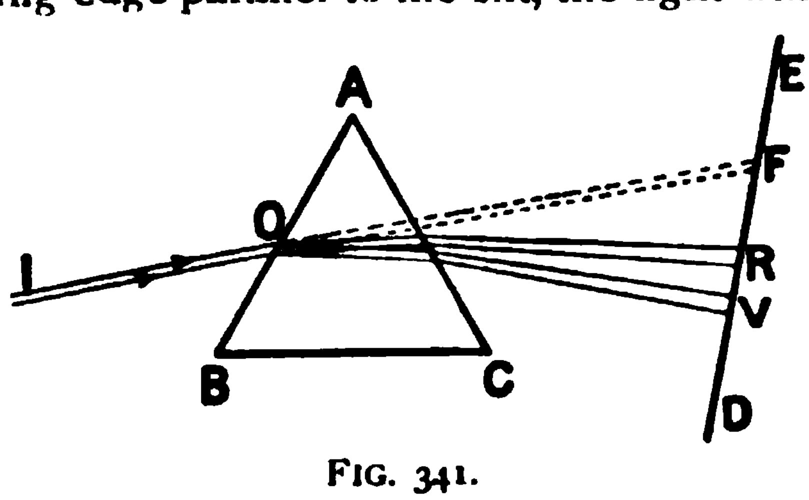


FIG. 341.

Thus white light has been split up by the prism into light of a number of different colours, these coloured lights being deviated to a different amount by the prism, so that the refractive index between two media, on which the deviation depends, is different for light of different colours; and since the violet rays are more deviated than the red, the refractive index for violet light is greater than for red light.

That white light is really formed by the superposition of light of all the colours of the spectrum can be shown by receiving the colours of the spectrum on a number of separate mirrors, and reflecting the light from them to the same point, when it will be found that white light will be reproduced.

In the form of the experiment described above, the different colours overlap on the screen to a certain extent; and in order to obtain a spectrum where no overlapping takes place, or a *pure spectrum*, as it is called, we may adopt the arrangement shown in Fig. 342. Light from a source L, such as the electric arc, passes through a narrow slit in a screen S, and then falls on a convex lens A, which, when the prism is not interposed, forms a real image of the slit at S′. If now the prism is interposed at B, the light will be deviated towards the base of the prism, and a spectrum will be formed on a screen placed at D. If we suppose that the slit is illuminated by violet light only, then an image of the slit will be produced at V, while if red light is used the image will be at R. Hence the spectrum VR is composed of a series of images of the slit formed by differently coloured light. If the slit is very narrow, one image will overlap very little on the adjacent images, and a pure spectrum will be obtained. As the slit is widened the images will overlap more and more, till with a very wide slit we shall get a white patch in the centre of the spectrum where all the images overlap, with a red edge at one end and a violet edge at the other.

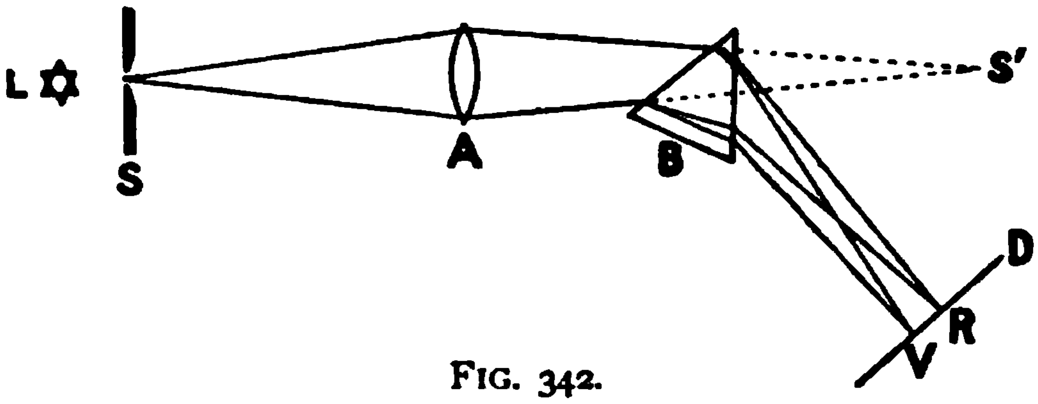


FIG. 342.

Another method of obtaining a pure spectrum is shown in Fig. 343. Parallel light, which may be obtained by means of a collimator, being

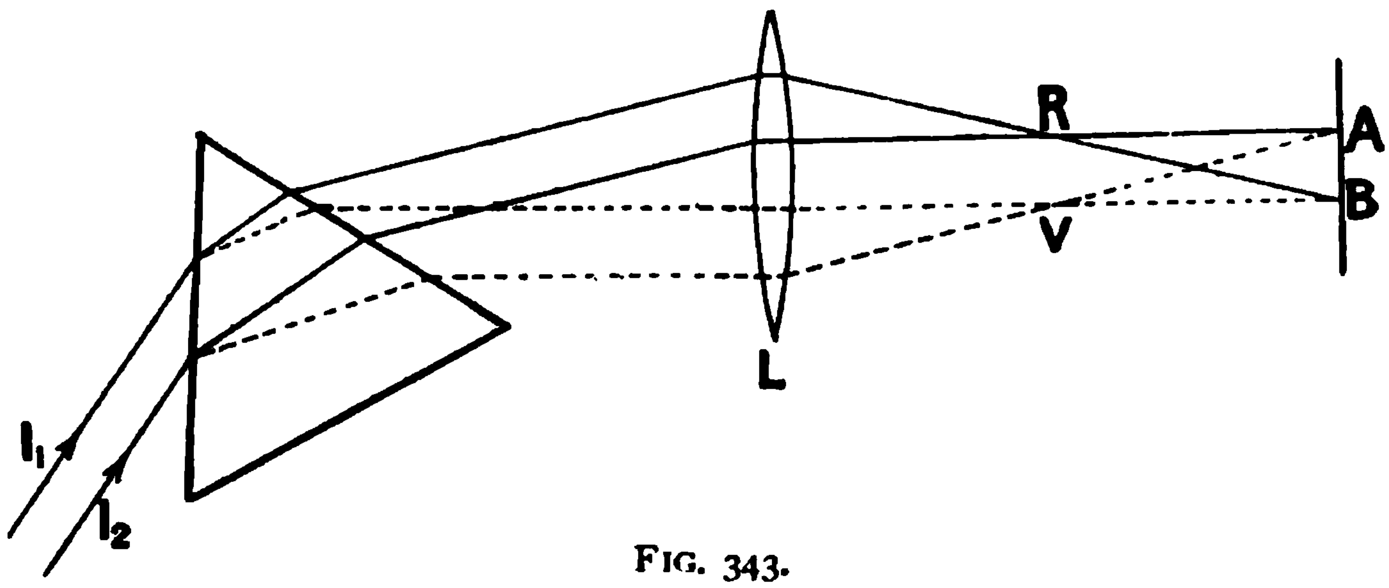


FIG. 343.

incident on the prism, a lens L is placed after the prism, and this lens brings the rays of the different colours to real foci between R and V, where a pure spectrum may be received on a screen or viewed with an eye-piece.

This arrangement will also allow of the recomposition of the different colours of the spectrum to form white, for if the screen be placed at AB the red and violet rays, as shown by the figure, and therefore also the rays of the other colours, will be uniformly spread over the patch AB. Under these circumstances a white patch will appear on the screen. If, however, a small obstacle be placed at V, so as to cut off the violet, the patch at AB will appear coloured a greenish-gold colour, produced by the mixture of the remaining colours. In the same way, by cutting off the red rays by an obstacle placed at R, the patch will appear a greenish blue.

When the slit of the spectrometer shown in Fig. 327 is illuminated with white light, a pure spectrum is formed at the principal focus of the lens F in the manner considered above, and can be observed with the eye-piece. The spectrometer, when used to observe spectra, is sometimes called a spectroscope. By using light of different colours, the refractive index of a substance for light of these colours can be obtained by any of the methods given in §§ 344, 346.

368. Fraunhofer's Lines.—When the slit of a spectroscope is illuminated by sunlight, it is found that the spectrum is traversed by an enormous number of dark lines parallel to the length of the slit. These dark lines are called Fraunhofer's lines, and are due, as we shall see later, to the light of the colours which are thus missing from the solar spectrum being absorbed in the sun's or the earth's atmosphere.

These lines form a very convenient means of specifying any particular colour in the spectrum, and hence the more prominent of them are indicated by the letters A, B, C, D, &c. Their relative position in the spectrum are shown in Fig 344. The lines A, B, and C are in the red, D

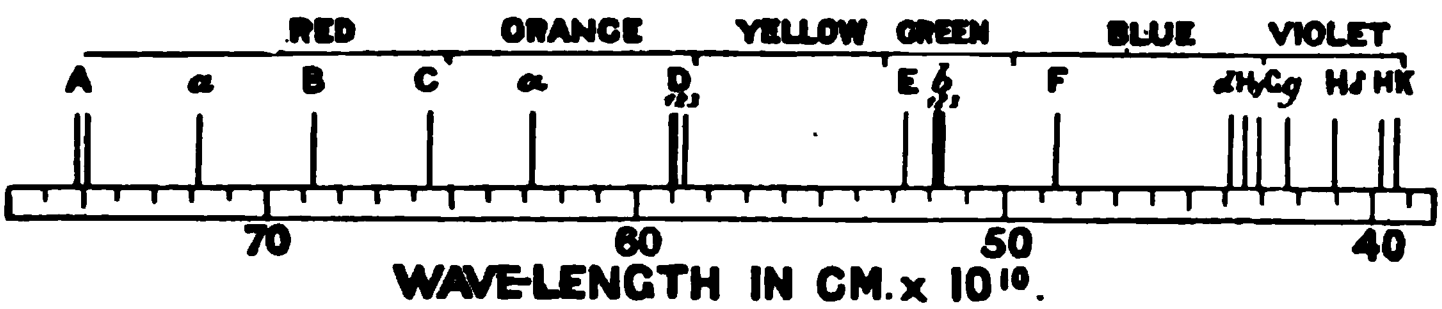


FIG. 344.

in the orange-yellow, E in the green, F in the greenish-blue, G in the indigo, and H in the violet part of the spectrum. Hence when we refer to light of any particular colour as, say, D light, we mean light of the colour which corresponds to the dark line D in the orange-yellow of the solar spectrum.

369. Refractive Index for Light of Different Colours—Dispersive Power.—In the following table the refractive index of some substances are given for the light corresponding to Fraunhofer's lines :—

REFRACTIVE INDEX.

	A.	B.	C.	D.	E.	F.	G.	H.
Water (16° C.)	1.330	...	1.332	1.334	...	1.338	...	1.344
Carbon bisulphide (10°)	1.616	...	1.626	1.635	...	1.661	...	1.708
Crown glass	1.528	1.530	1.531	1.534	1.537	1.540	1.546	1.551
Flint glass	1.578	1.581	1.583	1.587	1.592	1.597	1.606	1.614
Rock salt (17°)	1.537	1.539	1.540	1.544	1.549	1.553	1.561	1.568

When light passes through a prism the different colours are deviated to different degrees, so that if we have a parallel beam of light incident on the prism the rays of the different colours after passing through the prism will be inclined to one another. The angle between the emergent rays for any two colours is called the dispersion of these two colours, produced by the prism.

Since the rays are not only dispersed, but also deviated, it becomes of interest to see if there is a fixed relation between the dispersion and the deviation. The ratio of the dispersion for any two colours to the deviation of the mean ray between the two is called the dispersive power of the substance of which the prism is made.

We have seen, in § 345, that if the prism is at minimum deviation, then

$$\mu = \frac{\sin \frac{1}{2}(\delta + \theta)}{\sin \frac{1}{2}\theta}.$$

If the prism has a very small refracting angle, so that θ is very small, the deviation δ will also be very small, and hence the ratio of the angles $\theta + \delta$ and θ will be the same as the ratio of the sines. Thus for a prism of very small refracting angle

$$\mu = \frac{\theta + \delta}{\theta},$$

$$\therefore \qquad \delta = \theta(\mu - 1).$$

If μ_A is the refractive index, and δ_A the deviation for light corresponding to the Fraunhofer line A in the extreme red, μ_H and δ_H the corresponding quantities for the H line, and μ_D and δ_D for the D line, which may be taken as the mean light between A and H, we have

$$\delta_A = \theta(\mu_A - 1)$$
$$\delta_H = \theta(\mu_H - 1)$$
$$\delta_D = \theta(\mu_D - 1).$$

Hence the dispersion between the A and the H light is

$$\delta_H - \delta_A = (\mu_H - \mu_A)\theta.$$

Thus, if we take prisms having the same refracting angle θ, we get the following values for the dispersion :—

Water	$.014\ \theta$
Carbon bisulphide	$.092\ \theta$
Crown glass	$.023\ \theta$
Flint glass	$.036\ \theta$
Rock salt	$.031\ \theta$

Hence it will be seen that the extent of the spectrum obtained with a prism filled with carbon bisulphide will be 6.5 times as great as the spectrum obtained under similar circumstances with water.

The dispersive power is given by

$$\frac{\delta_H - \delta_A}{\delta_D} = \frac{\mu_H - \mu_A}{\mu_D - 1},$$

so that the dispersive powers of the substances given in the table above are :—

Water	0.042
Carbon bisulphide	0.145
Crown glass	0.043
Flint glass	0.061
Rock salt	0.057

370. Achromatic Prisms and Direct-Vision Spectroscopes.— Since the dispersive powers of different substances are not the same, we can obtain two prisms constructed of different materials such that, while the dispersion they produce is the same, the deviation produced on the mean ray is different, or *vice versa*. For instance, taking crown and flint glass, if we have a prism of flint glass of which the angle θ is small, so that we may apply the formula $\delta = \theta(\mu - 1)$, the dispersion is $0.036\ \theta$. If, then, we take a prism of crown glass of which the angle is ϕ, the dispersion will be $0.023\ \phi$. If the dispersion is to be the same in the two cases, we must have

$$\phi = \frac{0.036}{0.023}\theta = 1.56\ \theta.$$

Now the deviation produced by the flint-glass prism for the D line will be $(\mu_D - 1)\theta = 0.587\ \theta$, and the deviation for the same light produced by the crown-glass prism will be $(\mu'_D - 1)\phi = 0.534 \times 1.56\ \theta = 0.833\ \theta$. Hence, although the two prisms produce the same dispersion, the deviation produced by the crown-glass prism is greater than that produced by the flint-glass prism.

If the prisms are placed with their refracting edges turned in opposite directions, a ray of white light will, during its passage through the crown-glass prism C (Fig. 345), be deviated towards the base, *i.e.* downwards in the figure, and also dispersed, the red ray being less deviated than the

violet. During the passage of these rays through the flint-glass prism, however, the deviations of both rays will be in the opposite direction, *i.e.* upwards, the red ray being deviated upwards less than the violet. The difference between the deviations of the red and violet rays being the same in the two prisms, the rays when they leave the flint-glass prism will be parallel. The mean deviation in the crown prism being, however, greater than that in the flint, the rays on the whole will be deviated downwards, *i.e.* towards the base of the crown-glass prism.

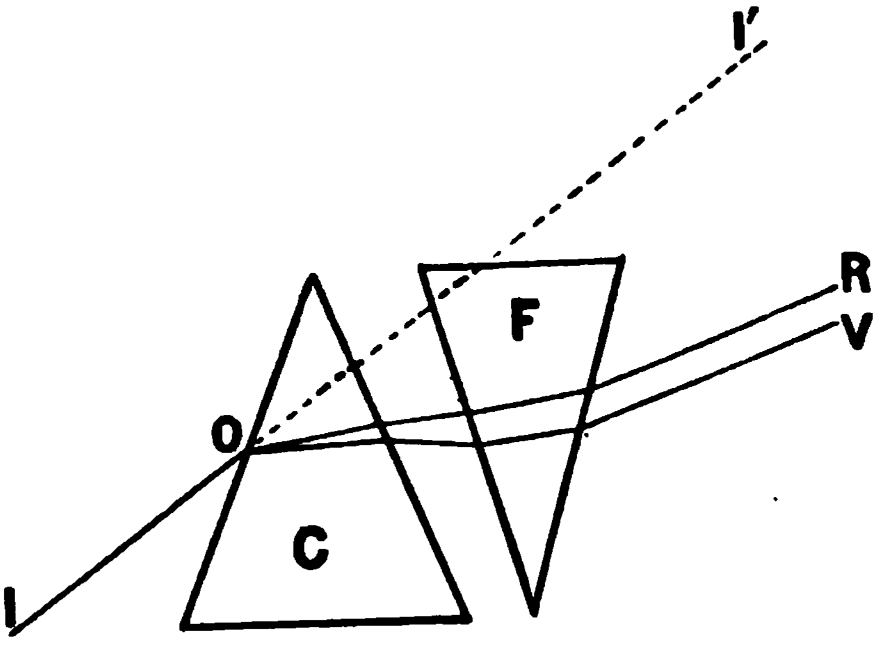


FIG. 345.

If we consider a second ray incident parallel to IO and close to it, it also will be split up, and the red and violet rays, after passing through the two prisms, will be parallel to the other two. A consideration of Fig. 346 will show that the patch of light received on a screen will be

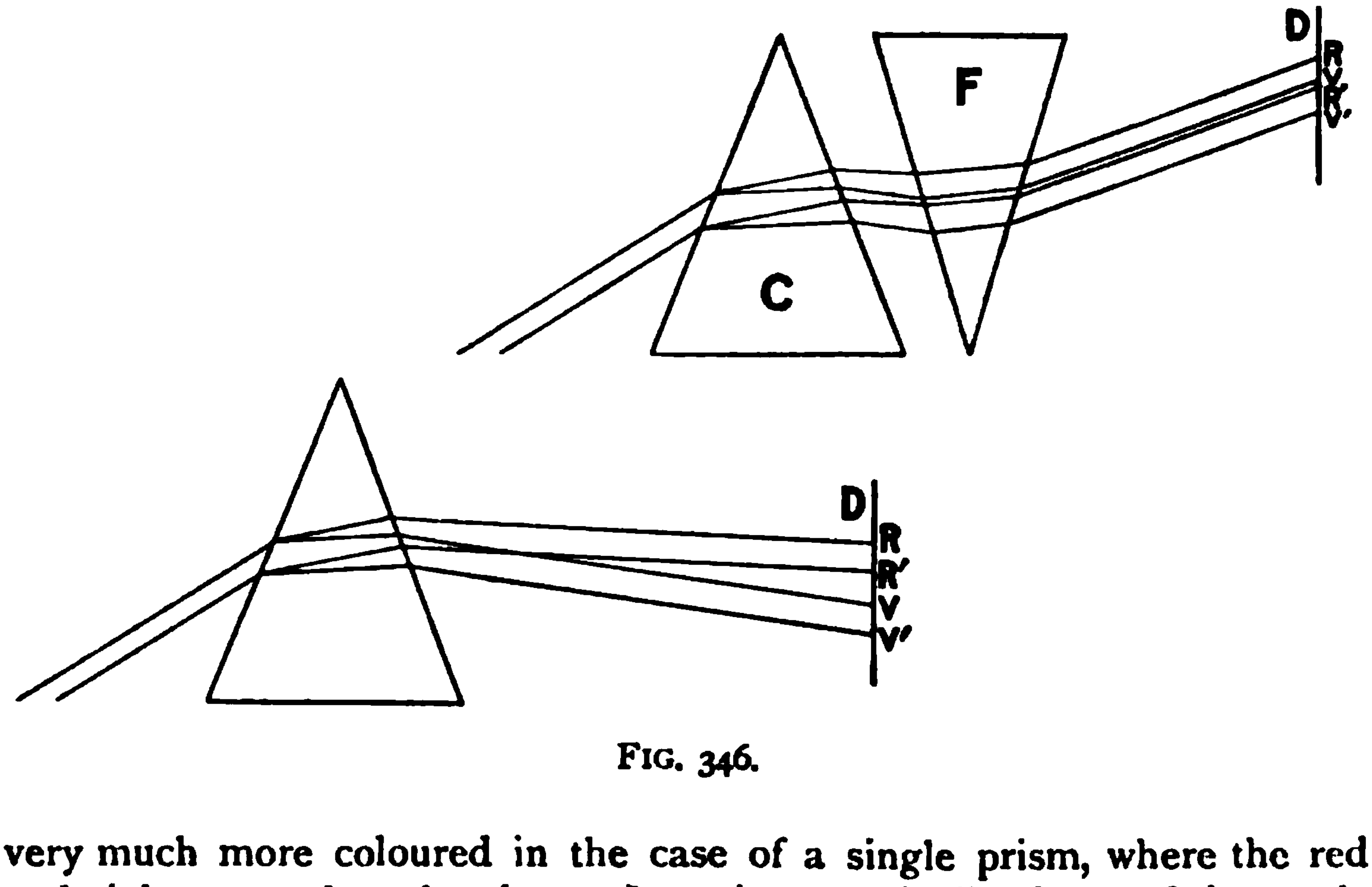


FIG. 346.

very much more coloured in the case of a single prism, where the red and violet rays when they leave the prism are inclined at a finite angle to one another, than in the case of the two prisms considered above, where the red and violet rays are parallel after they leave the second prism. For, considering only the two incident rays shown in the figure,

in the case of the single prism we shall have a white patch R′V, bordered by a wide red margin RR′ on one side, and a violet one VV′ on the other. In the case of the two prisms, however, the coloured margin is very narrow, and is of the same width at all distances from the prisms.

Hence by combining two prisms, one of crown glass and the other of flint, the ratio between the refracting angles having been suitably chosen, we are able to get a compound prism which deviates light but does not disperse it. Such a combination is said to be *achromatic.*

Instead of choosing the angles of the prisms such that the dispersion is the same, we might have chosen them such that the deviation produced on the mean ray was the same. In this case we have

$$\theta(\mu_D - 1) = \phi(\mu'_D - 1)$$
$$0.587\ \theta = 0.534\ \phi$$
$$\phi = 1.1\ \theta.$$

The dispersion produced by the flint-glass prism will be

$$\theta(\mu_H - \mu_A) = 0.036\ \theta,$$

and that produced by the crown-glass prism will be

$$\phi(\mu'_H - \mu'_A) = 0.023\ \phi = 0.023 \times 1.1\ \theta = 0.025\ \theta.$$

Hence the dispersion produced by the flint-glass prism is greater than that produced by the crown-glass one, so that if the prisms are placed with their refracting angles turned in opposite directions, the mean ray D will be undeviated by its passage through the two prisms, while the violet rays will be deviated one way and the red rays the other. In this way a spectrum is produced without the mean ray being deviated. This arrangement is used in some pocket forms of spectroscopes, which are called direct-vision spectroscopes, since one looks straight through the prisms at the slit, and not at an angle as in the ordinary form of spectroscope.

371. Achromatic Lenses.—In considering the formation of images by lenses, we have supposed that the light was monochromatic. When white light is used, we shall not only get the deviation which we have hitherto considered, but also dispersion.

Suppose we have a convex lens AB (Fig. 347), and that a parallel beam of white light falls on it, then where the rays enter and leave the lens the violet rays will be more deviated towards the axis of the lens than are the red rays, and hence the violet rays will be brought to a focus at a point V nearer the lens than the point R, where the red rays are brought to a focus. Thus if a screen is placed at V we shall get a central violet spot surrounded by a red ring, while if the screen is placed at R there will be a central red spot surrounded by a violet ring.

If a convergent pencil of rays is incident on a concave lens, CD, the violet rays are brought to a focus at a point V, which is *further* from the lens than the point R, where the red rays come to a focus.

If the convex and concave lens are of the same material and of equal focal length, on being placed close together the two dispersions counteract each other, but in this case the deviation would also be nil, and the whole would simply act like a plane slab. By making the convex lens of crown glass, and the concave lens of flint glass, we are, however, able, as in the case of the prisms, to obtain equal and opposite dispersion, and still have deviation in the direction of that produced by the crown glass, *i.e.* the combination will be a convex lens, so that it is possible to construct achromatic lenses.

By the use of two lenses it is possible to make a lens which shall be achromatic as far as light of any two colours is concerned. The combination will not, however, be achromatic for light of other colours.

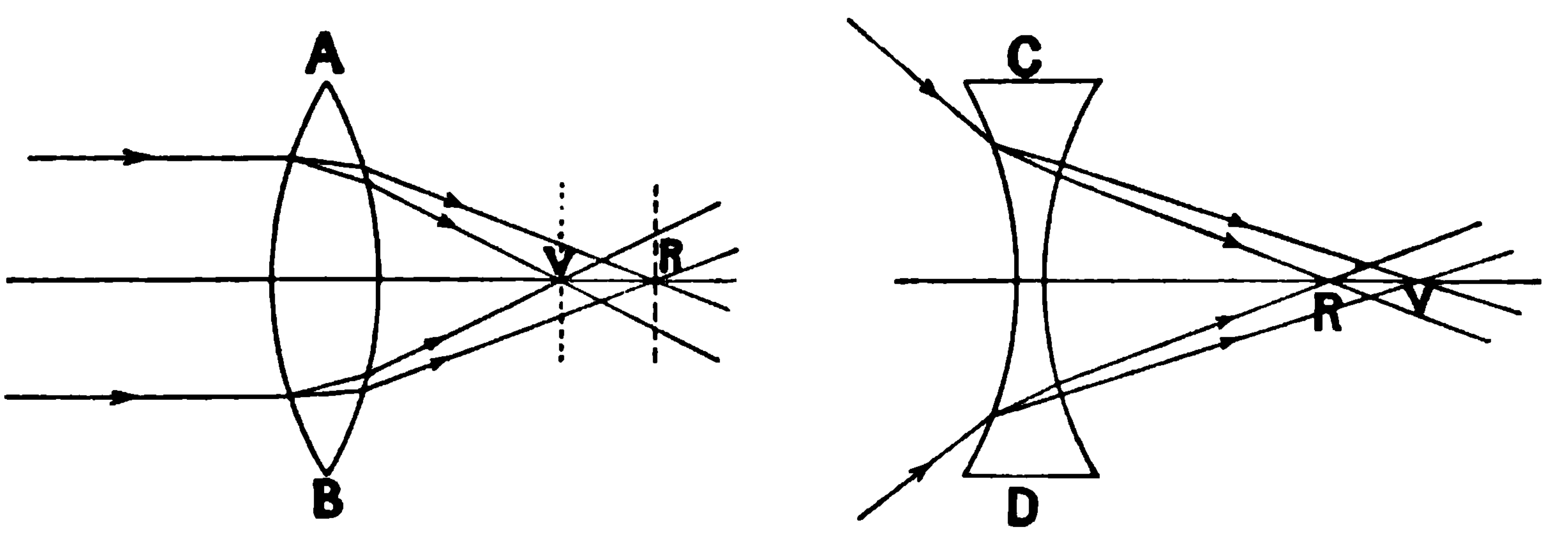


FIG. 347.

If in place of two lenses we use three, made of materials having different dispersive powers, the combination can be made achromatic for light of three colours, and so on. The colours for which the lens system is rendered achromatic depend on the purpose for which the lens is to be used. Thus for a telescope used in eye observations the colours chosen are those parts of the spectrum which affect the eye most strongly, while if the telescope is to be used for photography, it is most important that the lens should be achromatised for the violet and ultra-violet, since these rays are chiefly concerned with the production of the photographic image.

372. The Rainbow.—Let the circle in Fig. 348 represent a section of a spherical raindrop. When a ray of sunlight S_1M_1, which may be taken as parallel light, falls on the drop it will be refracted along M_1R_1, and when it comes to the surface at R_1, part of the light will be reflected along R_1N_1. On again reaching the surface at N_1 part will leave the drop, being refracted along N_1P_1. In the figure the paths of only a few of the rays have been drawn, in order to prevent confusion, but if all had

been drawn it would have been found that the rays incident near the point M_2, such that the radius OM_2 makes an angle of 59° with the direction of the incident light, are less deviated by their two refractions and one reflection than light incident at any other point. The figure also shows that the rays incident in the neighbourhood of M_2 form a parallel pencil, N_2P_2, when they leave the drop, while in the case of rays incident at any other point, M_3, they form a divergent pencil, N_3P_3. Now when we are dealing with a parallel pencil of rays, since the cross section remains constant, the decrease of the intensity of the illumination with the

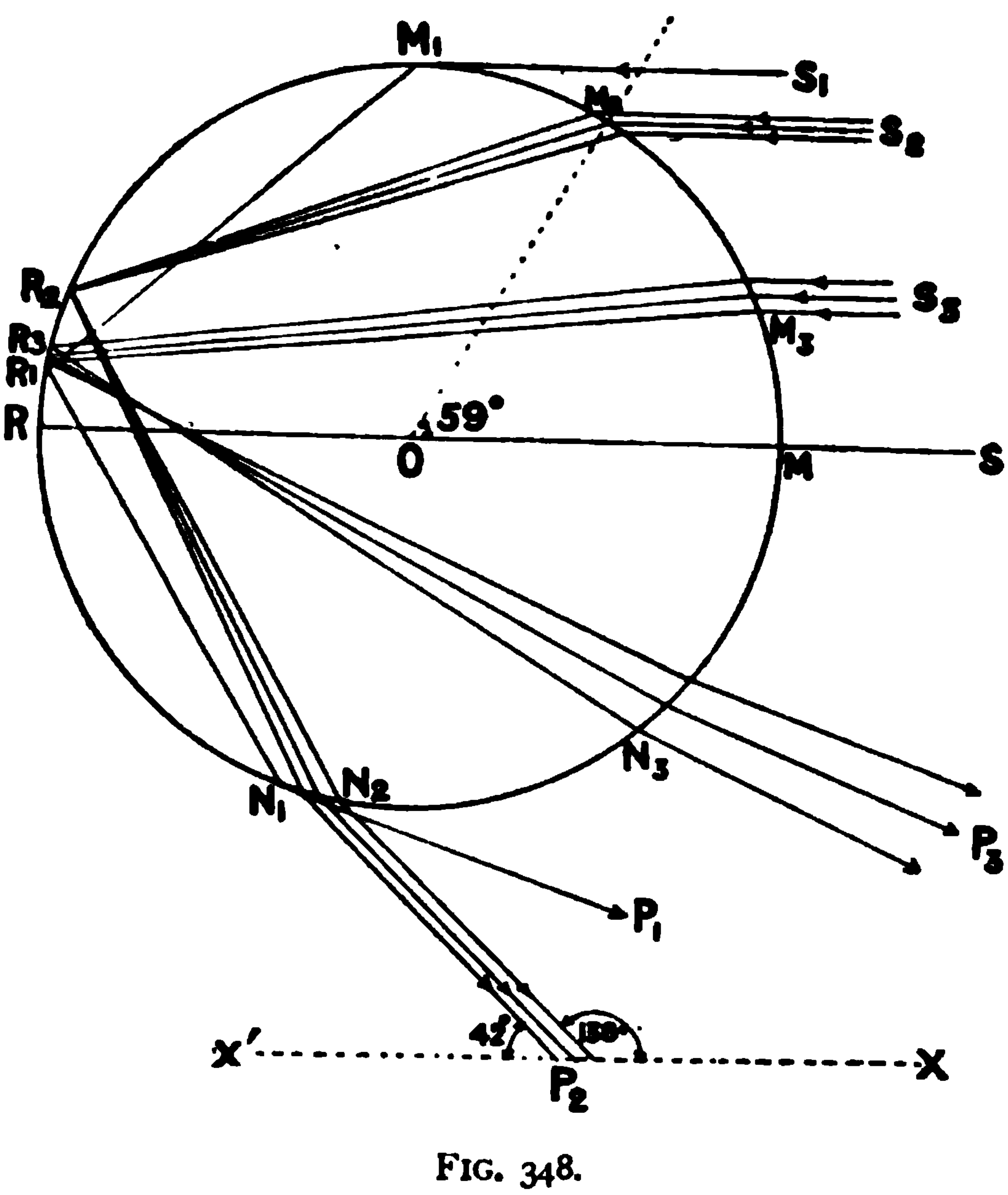


FIG. 348.

distance is small, being only due to absorption in the medium. With a divergent pencil it is, however, otherwise, for the rays are spread over a greater and greater area as we go away from their point of intersection, and hence the illumination decreases. Thus if we viewed such a raindrop from a distance, we should receive a considerable amount of refracted and reflected light if we looked along P_2N_2, but very little if we looked towards the drop in any other direction. If XX' is drawn parallel to the incident light, the angle N_2P_2X is 138°, and the angle N_2P_2X' is 42°. Hence supposing, with our eye, E (Fig. 349), as apex, and the direction of the sun's rays, SES', as axis, we describe a cone, of which the angle

between the generating lines and the axis is 42°, all raindrops, P_1, P_2, P_3, &c., which are on the surface of this cone will be so situated that the pencil of parallel rays which has undergone minimum deviation can enter the eye, and so the drops would be visible as bright points of light.

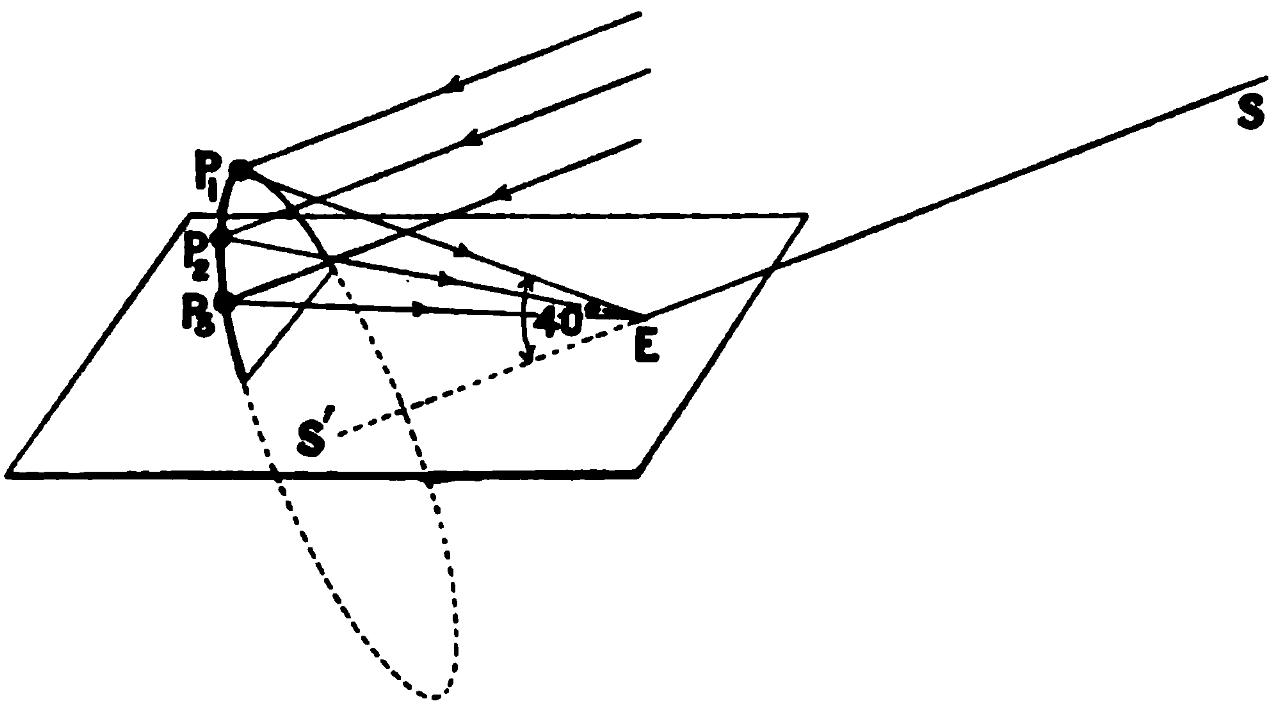


FIG. 349.

The phenomenon is not quite as simple as we have hitherto supposed, for the white sunlight is not only refracted when it enters and leaves the drop, but dispersion also takes place, as shown at A, Fig. 350. The result is that while the angle between the pencil of red rays which emerges

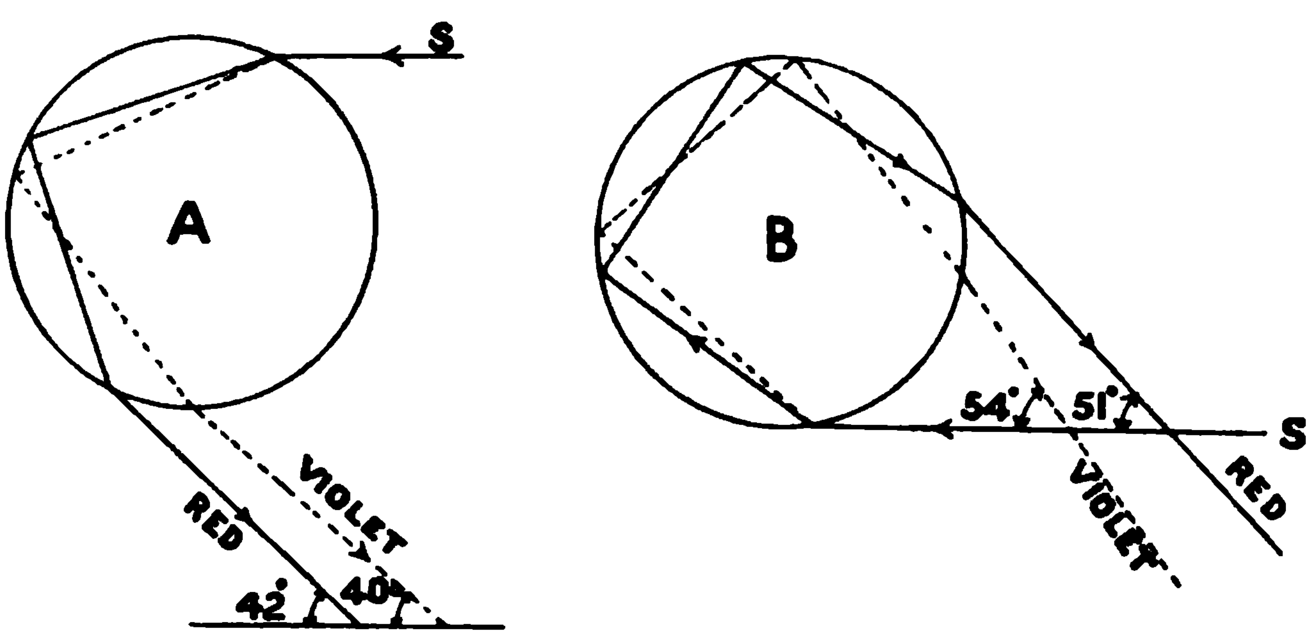


FIG. 350.

parallel and the incident light is about 42°, that between the violet rays is about 40°. Hence if we require to find the positions of the drops which will send violet light to the eye, we must construct a cone of which the half-vertical angle is 40°, which will of course lie inside the cone for the red rays. The cones corresponding to light of intermediate

wave-lengths will lie between these two, and therefore what is seen is a series of circular arcs showing the spectrum colours, the red being outside, and the other colours following in the order of descending wave-length, the whole constituting what is called the rainbow.

In addition to the bow which has been considered above, and which is called the primary bow, a secondary bow is sometimes seen outside the first. This bow is formed by light which has been twice reflected inside the raindrops in the manner shown at B (Fig. 350), and the angles of minimum deviation are 54° for the violet, and 51° for the red. In this bow, therefore, the violet appears on the outside.

CHAPTER VII

INTERFERENCE

378. Interference of Light.—The great difficulty met with at the outset by the exponents of the undulatory theory of light was the explanation of the rectilinear propagation of light and of the formation of shadows. In the case of the transmission of sound through air, which was admittedly due to the vibrations of the air particles, a sound produced outside a room, and coming in through the doorway, is found to spread all over the room, and does not confine itself to a beam passing across the room, as would be the case with light. It was only when the principle of interference was introduced into optics that the formation of shadows could be explained on the undulatory theory.

We have already seen how in the case of ripples on the surface of mercury, and in the case of sound, two wave-motions may combine together, so that while in some places they destroy each other, in others they strengthen each other.

In the case of light, all attempts to obtain interference between light waves emitted by two neighbouring sources, or even two separate portions of the same source, fails, this failure being due to the fact that the phase of the light vibrations given out by a source suffers rapid and abrupt changes, so that in the case of two separate sources the phase of the emitted light may be the same for, say, a thousand vibrations, a crest leaving each simultaneously, and thus producing darkness at a certain point P; then suddenly the phase of the light given by one source will change, so that while a crest is leaving one source a trough will be leaving the other, and thus the waves now strengthen each other at P. Since such changes, if they occurred a hundred times a second, would not be visible owing to persistence of vision, and during a hundredth of a second 5×10^{12} vibrations of yellow light take place, it is evident that it is not necessary to assume that the changes of phase take place so very frequently, in order to explain the absence of interference between the light from two independent sources.

If instead of two separate sources we take as sources two images of the same portion of a luminous body, then, whenever a change in phase takes place in the source, the corresponding change in phase will take place *simultaneously* on the two images; and hence if interference is produced at a given point before the change, it will also be produced after.

Fresnel, who first succeeded in demonstrating the interference of light, devised two arrangements for producing two image sources in such positions that they interfered. One of his arrangements consists of two mirrors, AB and BC (Fig. 351), inclined at an angle of very nearly 180°, so that a luminous point at P will produce two images, one at P′ by reflection in the mirror AB, and the other at P″ by reflection in the mirror BC. In § 331 we have proved that P′, P″ both lie on a circle of which B is the centre. Hence if we join P′P″, bisect this line at F, and join FB, FB will be at right angles to P′P″. If FB is produced to meet a screen on which the reflected light is received at O, then the point O is equidistant from P′ and P″.

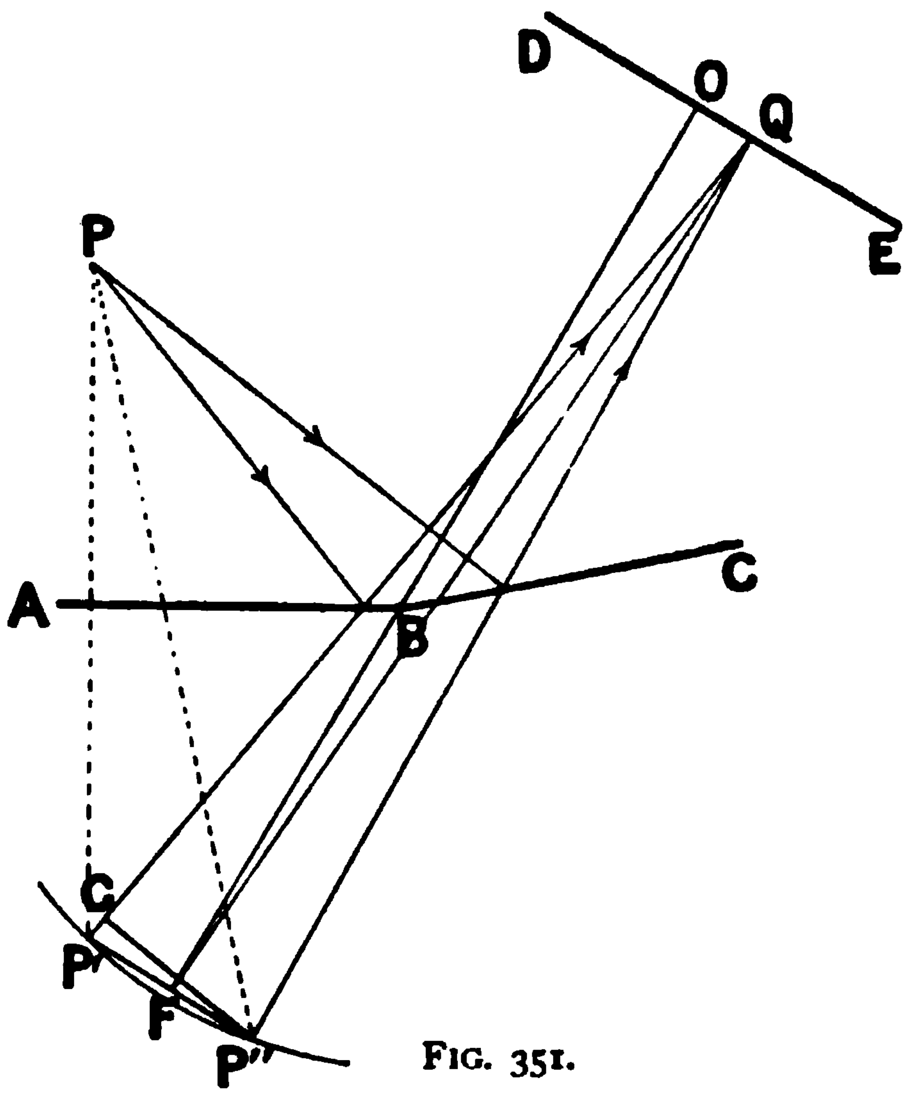


FIG. 351.

As far as the reflected light is concerned, we may regard it as coming from the images P′ and P″, so that the length of the path of any ray which, leaving P, is reflected at one of these mirrors and strikes a screen DE is the same as if it came from P′ or P″, as the case may be. At the point O of the screen, which is equidistant from the images, the light-waves will assist one another, since they always leave P′ and P″ in the same phase, these points being images of the same source.

There will be interference at a point such as Q, if the difference between the distances P′Q and P″Q is equal to half a wave-length, for then the vibrations from P′ will reach Q in the opposite phase to those from P″.

The manner in which the waves coming from P′ and P″ at some points on the screen strengthen each other, and at other points annul each other, is made clear by the diagrammatic representation given in Fig. 352. In this figure the waved lines represent the displacements proceeding from the two sources P′, P″ along the lines P′O, P″O, P′Q_1, P″Q_1, &c.; and it will be seen that the displacements produced at the points O and Q_2, due to the two sets of waves, are in the same phase, so that the resultant displacement is twice the displacement due to either.

It will also be seen that there are an equal number of waves between P′ and O and P″ and O, but that there is one more wave between P′ and Q_2 than between P″ and Q_2. At the point Q_1 the waves from the two sources are in opposite phase and destroy each other, and here there is half a wave more between P′ and Q_1 than between P″ and Q_1.

If with centre Q (Fig. 351) and radius QP″ we describe an arc of a circle cutting QP′ in G, then GP′ will be the difference between the paths P′Q and P″Q.

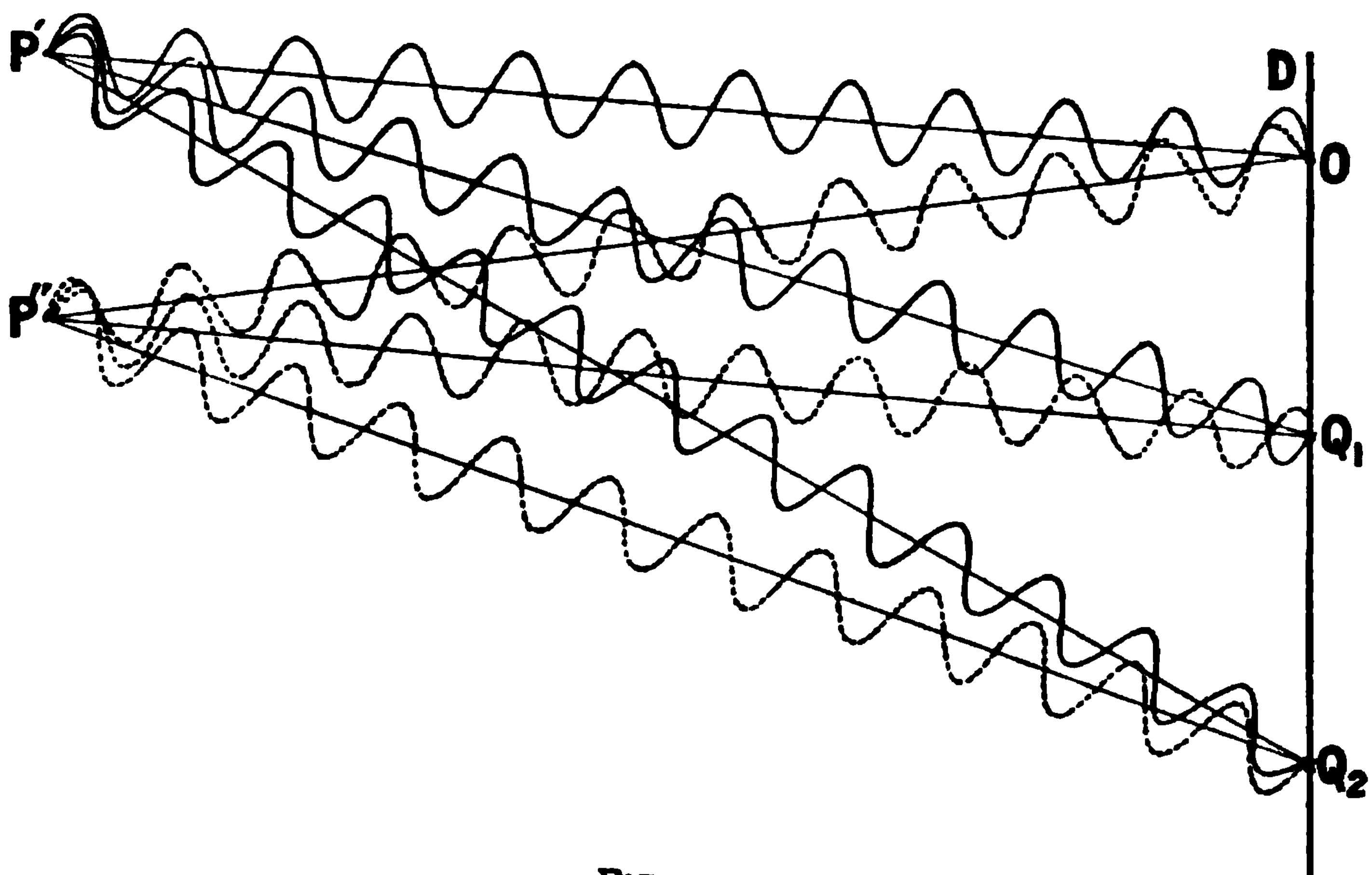


FIG. 352.

Since in practice the distance between the images P′ and P″ is excessively small compared with the distance of either of them from the screen, the arc P″G may be taken as a straight line, which is practically perpendicular to QP′ and QF.

Since P′P″ is perpendicular to OF, and GP″ is perpendicular to QF, the angle OFQ is equal to the angle GP″P′. Hence calling this angle θ, and the distance between the images $2d$, we have—

$$\overline{GP'} = 2d \sin \theta$$
$$= 2d \, . \, \frac{\overline{OQ}}{\overline{QF}}$$

or, since $\overline{QF}$ is practically equal to $\overline{OF}$, θ being very small,

$$GP' = 2d \, . \, \frac{\overline{OQ}}{\overline{OF}}.$$

Calling the distance BF between the images and the mirrors p, that

between the mirrors and the screen q, and the distance of the point Q from O x, we get, since OF $=p+q$,

$$\overline{GP}=\frac{2dx}{p+q}.$$

Now if λ is the wave-length of the light given out by the source P, interference will be produced at Q whenever the difference between the paths P′Q, P″Q is such that the light reflected from the two mirrors arrives at Q in opposite phases. This will occur when the paths differ by any odd number of half wave-lengths, for the ether at points separated by half a wave-length or any odd number of half wave-lengths is vibrating in opposite phases.

Hence if $\overline{GP'}$ is equal to $(2n+1)\frac{\lambda}{2}$, we shall get interference at Q.

Hence if there is interference

$$\frac{2dx}{p+q}=(2n+1)\frac{\lambda}{2},$$

or

$$x=(2n+1)\frac{\lambda}{2}\cdot\frac{p+q}{2d}.$$

If the difference in path, $\overline{GP'}$, is equal to an even number of half wave-lengths, *i.e.* to a whole number of wave-lengths, the light will reach Q in the same phase, and hence a bright band will be produced at Q. When this occurs,

$$x=2n\cdot\frac{\lambda}{2}\cdot\frac{p+q}{2d}.$$

In this experiment the two image sources P′, P″ play the same part as the two needle-points attached to the tuning-fork in the interference experiment with capillary waves on the surface of mercury (§ 271). If a line in Fig. 223 were drawn parallel to the line joining the two centres of disturbance, this would represent the screen in the optical experiment, and wherever this line cuts one of the interference curves in the figure would correspond to a dark band in the optical experiment, while half-way between each curve the mercury surface is disturbed by the combined action of the two centres of disturbance, and this corresponds to a bright band in light.

In the second method used by Fresnel, the light from a luminous point P passes through two narrow-angle prisms, AB (Fig. 353), placed with their bases in contact, forming what is called Fresnel's bi-prism. After passing through the bi-prism, the light travels as if it came from the two points P′ and P″, and interference is produced on a screen placed at DE, as in the previous case. Calling, as before, the distance of the images from the bi-prism p, that of the screen from the bi-prism q, and

the distance between the images $2d$, then a dark band will be produced by interference at Q, if the distance $\overline{OQ}$ or x is such that

$$x=(2n+1)\frac{\lambda}{2}\frac{p+q}{2d},$$

as in the case of the two mirrors.

Fresnel made use of these experiments on interference to prove that the velocity of light in air is greater than in glass, and hence to show that the emission theory was untenable.

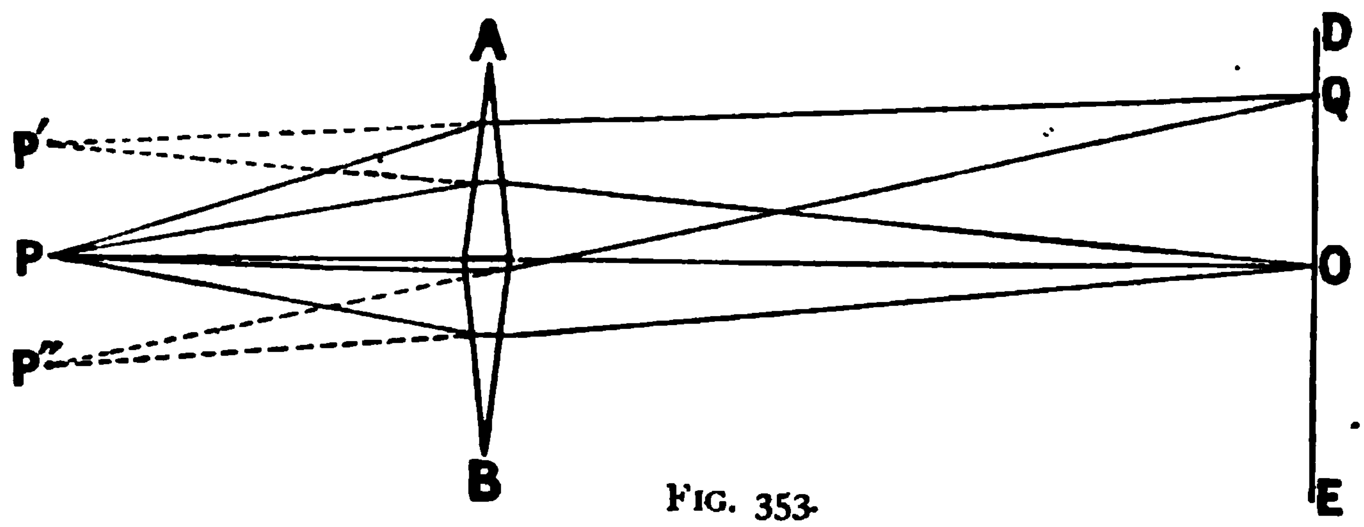


FIG. 353.

Let v_a be the velocity of light of any given colour in air, and v_g the velocity of the same coloured light in glass, and λ_a and λ_g the wave-length in air and glass respectively. The colour of the light being the same whether it is passing through air or glass, the frequency of the vibration must be the same in the two media, so that we have

$$v_a = n\lambda_a,$$

and

$$v_g = n\lambda_g,$$

or

$$\frac{v_a}{v_g}=\frac{\lambda_a}{\lambda_g}.$$

Suppose now that light proceeding from the two points P′ and P″ (Fig. 354) produces interference at Q, and that in the path of the light proceeding from P″ we introduce a plate of glass AB of thickness y. Further, suppose that originally Q was the first dark interference band, so that the paths P′Q and P″Q differed by $\lambda/2$. Before the introduction of the glass the number of wave-lengths in air between A and B was y/λ_a, while after the introduction of the glass the number of wave-lengths between A and B is y/λ_g. If then λ_g is less than λ_a, there will be more wave-lengths between A and B when the glass is introduced than there was before, so that the

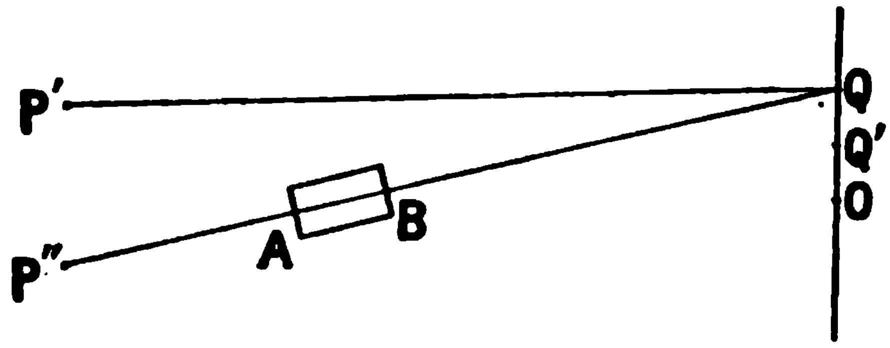


FIG. 354.

path from P″ to Q will be longer, *i.e.* contain a greater number of wave-lengths than before, and therefore the two paths P′Q and P″Q will no longer differ by half a wave-length. In order to get interference when the glass is introduced, we must therefore lengthen the path P′Q and shorten the path P″Q. This will be the case for a point such as Q′ nearer O than Q. In the same way, if λ_g is greater than λ_a, the point where interference is produced will be moved away from O when the glass is introduced, so that if, on the introduction of the glass, the interference bands move towards the side on which the glass has been introduced, we should infer that the wave-length in glass λ_g was smaller than the wave-length in air λ_a, and *vice versa*. On performing the experiment the bands move *towards* the side on which the glass is introduced, so that the wave-length in air, and hence also the velocity in air, is greater than the corresponding quantities for glass.

Fresnel's experiments may be used to measure the wave-length λ of the light used, for by measuring x we can calculate λ. It is, however, generally more convenient to measure the distance between the dark interference bands than to measure the distance of a dark band from the central bright band. The distances of the first few dark bands from O are given by

$$x_1=\frac{\lambda}{2}\cdot\frac{p+q}{2d},$$

$$x_2=\frac{3\lambda}{2}\,\frac{p+q}{2d},$$

$$x_3=\frac{5\lambda}{2}\,\frac{p+q}{2d}, \text{ \&c., \&c.}$$

so that the distance (z) between two consecutive bands is

$$z=\lambda\frac{p+q}{2d}.$$

In this expression $\frac{p+q}{2d}$ is independent of the wave-length of the light used, so that we see that the greater the wave-length λ, the further apart are the interference bands. It is found by experiment that with red light the bands are further apart than with violet light, so that the wave-length of violet light must be less than that of red light.

If white light is used, the violet light will be destroyed nearer to the centre O than the red light, so that this red left over will produce a red band on either side of the central bright band, which will be white, for the light of all wave-lengths arrives in the same phase at O. A little further out from O the red light will be destroyed by interference leaving the violet light, so that here a violet band will be produced. Hence when white light is used the first dark band will be bordered with red on the inside and violet on the outside. The distance between the points where the red and violet are destroyed will increase with each successive

band, until finally there will be overlapping between one bright band for the violet and the previous bright band for the red ; and at some distance from O, the overlapping will be so considerable that white light will be reproduced, and so no bands will be discernible.

374. The Diffraction Grating.—A diffraction grating consists of a number of equidistant parallel lines ruled on a plate of glass, or of speculum metal. In order to explain the action of a grating, we shall suppose that it consists of a series of equally spaced opaque lines ruled on a plate of glass, the width of each line being equal to the space between two adjacent lines.

Let AB and CD (Fig. 355) be two adjacent spaces, and suppose a beam of parallel light to be incident on the grating normally, *i.e.* parallel to NA, so that the incident wave-fronts are parallel to the grating. We may then look upon each point in the spaces AB, CD, &c., as a centre of disturbance from which light-waves are propagated, all these waves starting in the same phase. Consider two of these centres of disturbance, one at A and the other at C. The disturbances from these centres will reach all points at equal distances from A and B in the same phase, and so will strengthen one another.

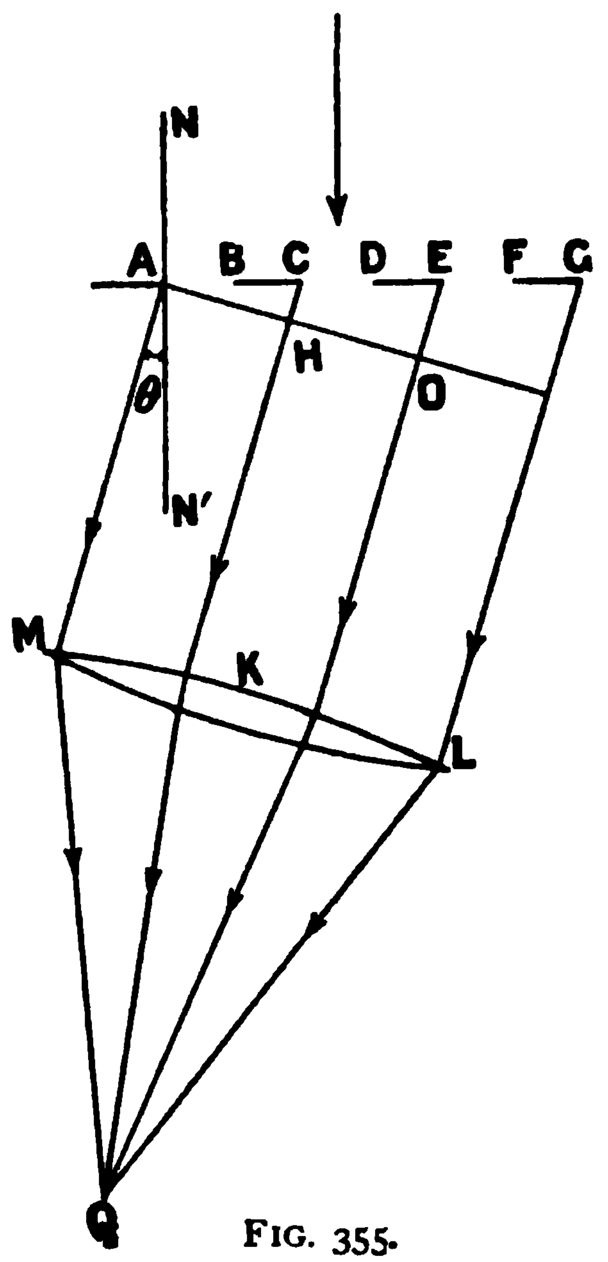


FIG. 355.

At any other point Q, however, the disturbances need not be in the same phase. If, as is always the case, Q is at very great distance from the grating compared to AB, or, as shown in the figure, a lens L is interposed to form an image at its principal focus, we may take the lines AM and CK as parallel, and both inclined to the normal to the grating at an angle θ.

From A draw AH perpendicular to CK or AM ; then, since N′A is perpendicular to CA, and AH is perpendicular to AM, the angle CAH included between CA and AH is equal to the angle θ included between N′A and AM. Therefore

$$\overline{CH}=\overline{AC}\sin\theta,$$

or if d is the combined width of a space and a line, so that AC$=d$,

$$\overline{CH}=d\sin\theta.$$

Now the waves starting from A and C will be in the same phase when

they reach Q, and therefore will strengthen each other, if the difference in the paths $\overline{AM}$ and $\overline{CK}$ is equal to an even number of half wave-lengths.[1] Hence the condition for the production of a bright band at Q is that

$$d \sin \theta = 2n . \frac{\lambda}{2},$$

or

$$\sin \theta = \frac{n\lambda}{d}.$$

If $\overline{CH}$ is equal to an odd number of half wave-lengths, interference will be produced at Q, the condition for a dark band being

$$d \sin \theta = (2n+1)\frac{\lambda}{2}.$$

What we have said with regard to the two centres A and C will also apply to each pair of centres taken in AB and CD, so that the above equations also give the conditions for the production of a bright or dark band at Q, when the whole of the two spaces AB and CD are operative. A similar argument holds with regard to the next two spaces, and so on, so that the above equations apply to the grating taken as a whole.

If we consider the first bright band, then, as d is constant and $\sin \theta = \lambda/d$, it is evident that the value of θ will vary with the wave-length of the light, so that by measuring the angle θ for the first bright band produced by different coloured lights, we can calculate the wave-length of these lights. If white light is used, the positions of the bright bands will be different for the different colours, and hence a spectrum will be produced.

When using this method to measure the wave-length of light the grating is mounted on the table of the spectrometer (Fig. 327), with its surface normal to the light coming through the collimator and the rulings on the grating parallel to the slit. The telescope is then turned to view the bright bands on either side of the central bright band corresponding to the undeviated light, and the difference between the readings gives 2θ.

If white light is used, a series of spectra will be obtained corresponding to the cases where n is made 1, 2, 3, &c., in the formula

$$\sin \theta = \frac{n\lambda}{d}.$$

The least deviated spectrum, for which $n = 1$, is called the first spectrum; and if λ_A, λ_H are the wave-lengths of the light corresponding to the

[1] After striking the lens the waves will be brought to a focus at Q, and the virtual length of the paths depends on the constants of the lens. The virtual length, that is, the length allowing for the fact that light travels slower in glass than in air, of all the paths from the first surface of the lens to the focus Q is the same, so that any difference of phase which exists at N and K will persist when the waves reach Q.

A and H Fraunhofer lines, the difference between the values of θ for these two lines will be given by

$$\sin\theta_{\rm A} - \sin\theta_{\rm H} = \frac{1}{d}(\lambda_{\rm A} - \lambda_{\rm H}).$$

In the second spectrum, for which $n=2$, we have in the same way—

$$\sin\theta'_{\rm A} - \sin\theta'_{\rm H} = \frac{2}{d}(\lambda_{\rm A} - \lambda_{\rm H}).$$

Hence in this second spectrum the difference between the sines of the angles of deviation is twice as great as in the first spectrum, so that $\theta'_{\rm A} - \theta'_{\rm H}$ is greater than $\theta_{\rm A} - \theta_{\rm H}$, or, in other words, the dispersion in the second spectrum is greater than in the first. In the same way, the dispersion in the third spectrum is greater than in the second, and so on.

The grating, when suitable precautions are taken, is a marvellously accurate means of measuring the wave-length of light, so that we are able to measure these extremely minute lengths to within about one part in 60,000.

In the following table the values of the wave-length for the principal Fraunhofer lines are given :—

WAVE-LENGTH OF FRAUNHOFER LINES IN AIR.

A . .	{ 7621 × 10^{-8} cm. / 7594	E . .	5271 × 10^{-8} cm.
		F . .	4861
B . .	6870	G . .	4308
C . .	6563	H . .	3969
D_1 . .	5896	K . .	3934
D_2 . .	5890	L . .	3821
D_3 . .	5876	M . .	3728

375. Colours of Thin Plates.—It is a matter of common observation that some bodies, such as soap-bubbles, thin films of oil on water, of oxides on metals, glass, &c., show under certain conditions of illumination brilliant colours. The explanation of these colours on the wave-theory of light is easily obtained.

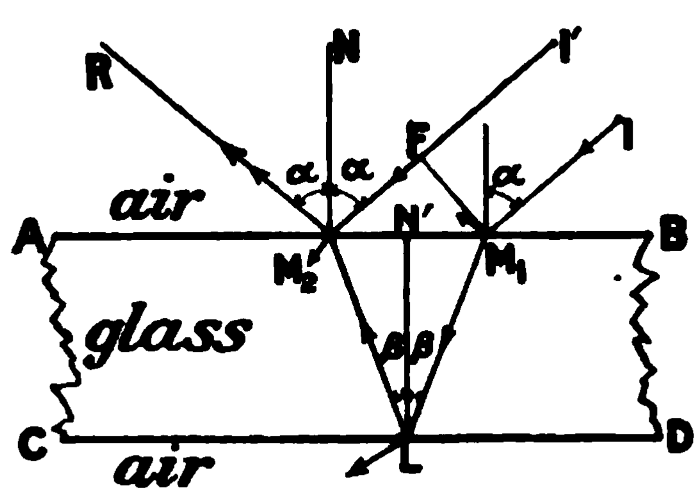


FIG. 356.

Let ABCD (Fig. 356) represent a section of a glass plate, and suppose that a parallel beam of light is incident in the direction IM_1. A ray, IM_1, will be partly reflected at M_1 and partly refracted along M_1L. At L the refracted ray will be partly reflected along LM_2 and partly refracted. The reflected portion will again meet the surface at M_2, where part will be reflected and part refracted

along M_2R. A ray directly incident at M_2 will also be partly reflected along M_2R, so that we shall have two waves which left I and I′ in the same phase, one having traversed the path $IM_1L.M_2R$ and the other the path $I'M_2R$, and as these paths differ in length, we may have these waves interfering.

In order to get the difference in these two paths, we draw M_1F perpendicular to $I'M_2$, and also the normals LN' and M_2N. Then $\overline{M_1L}+\overline{M_2L}=2\overline{M_1L}$, since the angle of reflection at L is equal to the angle of incidence. Hence if T is the thickness of the glass plate,

$$\overline{M_1L}+\overline{M_2L}=2\overline{M_1L}=2\frac{\overline{N'L}}{\cos\beta}=\frac{2T}{\cos\beta}.$$

Also $\quad FM_2=\overline{M_1M_2}\cos FM_2M_1=\overline{M_1M_2}\sin NM_2F=\overline{M_1M_2}\sin\alpha,$

and $$\overline{M_1M_2}=2\overline{M_2N'}=2T\tan\beta.$$

So then $$\overline{FM_2}=2T\sin\alpha\tan\beta.$$

The two paths not only differ in length, but also in that while FM_2 is in air, the path M_1LM_2 is in glass. Now the velocity of light in glass is to that in air in the ratio of 1 to μ, where μ is the refractive index from air to glass. Hence the effective length of the path $M'LM_2$ is $\mu(M_1L+LM_2)$ or $2T\mu/\cos\beta$. Thus the effective difference in the paths is

$$\frac{2T\mu}{\cos\beta}-2T\sin\alpha\tan\beta.$$

But $$\mu=\frac{\sin\alpha}{\sin\beta},$$

$$\therefore \qquad \sin\alpha=\mu\sin\beta.$$

Hence the effective difference in the paths is

$$\begin{aligned}&\frac{2T\mu}{\cos\beta}-2T\mu\sin\beta.\tan\beta\\&=\frac{2T\mu}{\cos\beta}-\frac{2T\mu\sin^2\beta}{\cos\beta}\\&=\frac{2T\mu}{\cos\beta}(1-\sin^2\beta)=\frac{2T\mu}{\cos\beta}\cos^2\beta\\&=2T\mu\cos\beta.\end{aligned}$$

If this difference in path is zero, then we should expect the two waves to strengthen each other, since along M_2R they would be in the same phase. If T is made vanishingly small the difference in the paths vanishes, so that there ought to be maximum reflection for this case, since the light reflected from the two surfaces will be in the same phase. It is found, however, in the case of a soap film, that as it gets thinner a thickness is at length reached when no light is reflected from the film, while when it

is thicker colours are produced. We are therefore led to the conclusion that the above investigation is defective. The fact is that the reflections that take place at L and M_2 occur under different conditions, in that at L the light is travelling in a dense medium and is reflected at a surface separating a denser medium (glass) from a less dense medium (air), while at M_2 the light is travelling in the less dense medium and is reflected at a surface separating this medium from a more dense medium.

We may form an idea in what manner this difference in the circumstances of the reflection will affect the phase of the reflected light by considering the impact of two elastic particles of different mass. If the lighter particle strikes the heavier particle, it will drive the heavier particle forward, but it will itself rebound so that its own motion will be suddenly reversed. Since a sudden reversal of a moving particle's motion corresponds to a change of phase of half a wave-length, we can conceive that when a light-wave moving in air strikes a denser medium, such as glass, the refracted ray will be in the same phase as the incident ray, but the reflected ray will undergo a sudden change in phase, equivalent to the loss of half a wave-length at the moment of reflection.

If a heavier particle strikes a lighter, then the lighter particle is driven forward, but the motion of the heavier particle continues in its original direction. So that, in the case of a wave of light travelling in a denser medium and meeting a surface separating this medium from a less dense one, there will be no change in phase in either the refracted or reflected wave.

Hence while the reflected ray at M_2 loses or gains, whichever we like to take it, half a wave-length, due simply to the reflection, the ray IM_1LM_2R does not suffer any such sudden loss or gain. In considering the interference of the two rays along M_2R, we must therefore add $\lambda/2$ to the path IM_2R, so that the difference in path, allowing for this effect due to reflection, is

$$2T\mu \cos \beta + \frac{\lambda}{2}.$$

If in this expression T is made very small, the two waves will differ in phase by $\lambda/2$ and hence will produce interference, and we shall get no reflected light, which agrees with experiment.

Interference will also take place if the difference in phase between the two rays is any odd number of half wave-lengths. Hence interference will take place if

$$(2n+1)\frac{\lambda}{2} = 2T\mu \cos \beta + \frac{\lambda}{2},$$

or if

$$n\lambda = 2T\mu \cos \beta.$$

In this expression it must be remembered that λ is the wave-length of the light in air, and β is the angle of incidence on the second surface of the thin plate.

If the thickness T of the plate varies and the incident light is white, a

series of coloured patches and streaks will be formed, for light of the different colours which constitute white light will be destroyed by interference at different points, the thickness being given by the above expression, and the reflected light will, by the loss of the destroyed rays, appear coloured.

376. Newton's Rings.—When a convex lens of large radius is pressed on a flat piece of glass or on a concave glass surface of greater radius of curvature than that of the lens, the point where the lens touches the glass will be seen surrounded by a series of dark rings if the light is monochromatic, or of coloured rings if white light is used. These rings, which are known as Newton's rings, are produced by interference in the thin film of air enclosed between the two glass surfaces, and may be seen both in the reflected and in the transmitted light.

If SOPN (Fig. 357) represents a section of the sphere from which the lens may be supposed to be cut, and AB the glass plate, which the lens touches at O, then the thickness of the air film included between the lens and the plate is zero at O and increases as we pass out from O.

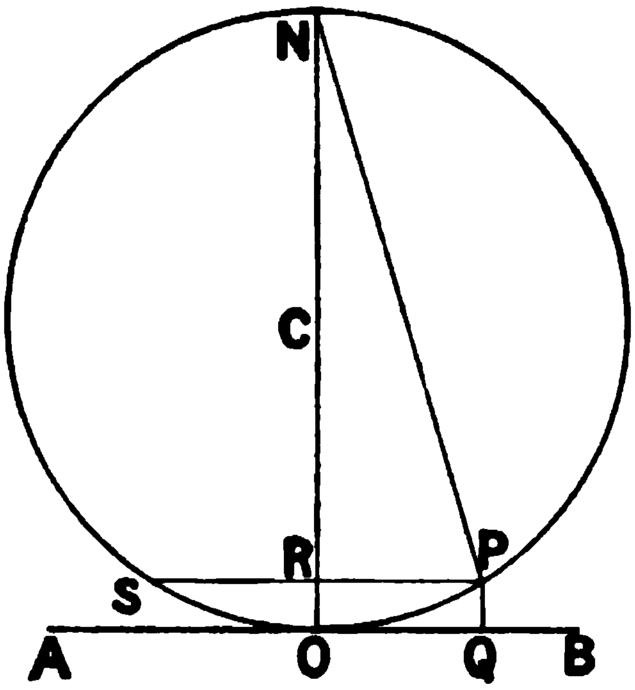


FIG. 357.

Let Q be a point at a distance r from the point of contact, then the thickness of the air film at Q can be found as follows. Draw QP perpendicular to AB to meet the circle, and through P draw PS parallel to AB, cutting the diameter, NO, of the circle in R. Then by a well-known property of the circle $\overline{PR}.\overline{RS}=\overline{OR}.\overline{RN}=(\overline{ON}-\overline{OR})\overline{OR}$.

Hence if R is the radius of curvature of the surface of the lens, and $\overline{OR}$ or $\overline{PQ}$, the thickness of the air film, is called T, we have

$$r^2=(2R-T)T=2RT-T^2.$$

Now since, when interference takes place, the thickness T of the air film is always very small compared to the radius of the lens R, the quantity T^2 will be very small compared to $2RT$, so that we may neglect T^2, and

$$r^2=2RT,$$

or

$$T=\frac{r^2}{2R}.$$

We have already seen that in the case of reflected light interference will take place if $T=\frac{n\lambda}{2\mu\cos\beta}$, where λ was the wave-length in the

medium outside the film. If λ_1 is the wave-length in the medium outside the film, and λ_2 the wave-length in the film, then

$$\frac{\lambda_1}{\lambda_2} = {}_1\mu_2,$$

or

$$\lambda_1 = \lambda_2\, {}_1\mu_2.$$

Hence interference will take place if

$$T = \frac{n\lambda_2}{2 \cos \beta},$$

where λ_2 is the wave-length in the film.

In the case of Newton's rings we are dealing with air as the film, and so λ_2 is here the wave-length in air, so that we shall have a dark ring passing through Q if

$$\frac{r_n^2}{2R} = \frac{n\lambda_2}{2 \cos \beta},$$

or if

$$r_n^2 = \frac{n\lambda_2 R}{\cos \beta}.$$

If the lens and plate are in contact at the centre, we shall get interference, as we have already shown, and there will be a black spot at the centre; the radii of successive dark rings will be obtained by taking n equal to 1, 2, 3, &c., so that the squares of the radii of successive dark rings are proportional to the natural numbers, 1, 2, 3, &c. The angle β is that which the light rays in the air film make with the plate, but in all practical cases the lens is of such small curvature, and the rings are only formed so near the centre, that we may regard the lens as a parallel plate, so that the rays in the air film will be parallel to the rays incident on the upper surface of the lens, and we may take β as the angle of incidence of the rays on the lens.

If the light is incident normally $\beta = o$, and

$$r_n^2 = n\lambda_2 R.$$

At the centre there is interference for all the colours, so that with white light the centre is black, as we pass out; if λ_H is the wave-length of violet light, then when r is equal to $\lambda_H R$ this violet light will be destroyed, and hence the remaining light will, along this circle, appear coloured red. A little further out, r is equal to $\lambda_A R$, so that the red light is destroyed, and the remaining light appears violet. When r is equal to $2\lambda_H R$, the violet will again be destroyed and the red left, while when r is equal to $2\lambda_A R$, the violet will be left. Thus with white light the central black spot will be surrounded by a series of coloured rings, each of which is red on the inside and violet on the outside.

Newton's rings are also formed in the light which is transmitted

through the lens and plate. If AB (Fig. 358) is the surface of the plate, and ADE the lens, then a ray of light incident normally along IQ traverses the air film, and at P is partly transmitted along PS and partly reflected along PQ′, where it is again partly reflected along Q′P′S′. Interference may then take place between the waves which have traversed the air film once, and those which have traversed it three times.

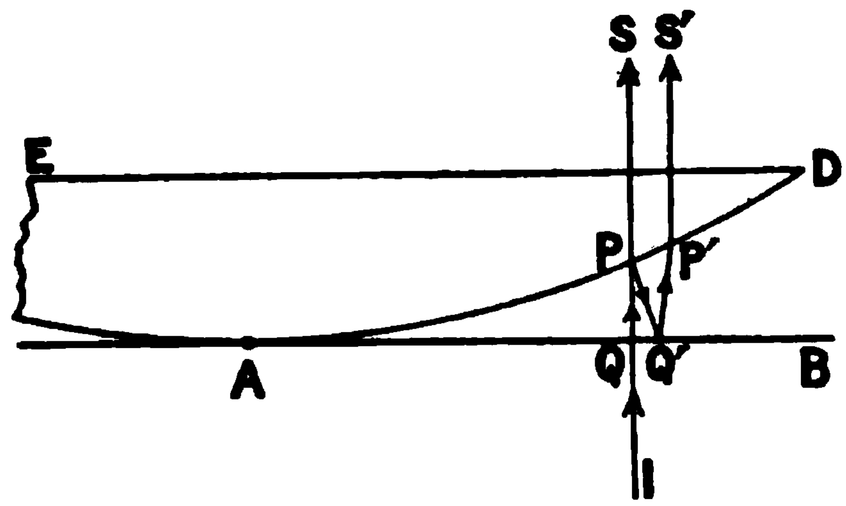


FIG. 358.

Using the same notation as before,

$$r^2 = 2R \,.\, \overline{PQ},$$

and the difference in path is given by

$$2PQ = \frac{r^2}{R}.$$

The ray IS, as it is nowhere reflected, undergoes no sudden change in phase ; the ray IPQ′S′, however, is reflected at P and at Q′, and in each case at the surface of a *denser* medium, and loses at each half a wave-length, or a whole wave-length in all. Hence, as the loss or gain of a whole wave-length by one ray does not affect the interference phenomena between two rays, we have that interference will take place when the difference in the paths is equal to an odd multiple of the half wave-length. Hence there will be a dark ring passing through Q if

$$2\overline{PQ} = (2n+1)\frac{\lambda}{2},$$

or

$$\frac{r^2}{R} = (2n+1)\frac{\lambda}{2},$$

or

$$r^2 = (2n+1)\frac{\lambda R}{2};$$

while there will be a bright ring for

$$r'^2 = 2n\frac{\lambda R}{2}.$$

When $n=0$, $r'=0$, so that there will be a bright spot at the centre, as is also obvious since here the lens and plate are in contact, so that there is no air film, and the light is simply transmitted. By comparing the expressions for r in the case of reflected and transmitted light, it will be seen that where there is a dark ring for one, there will be a bright ring for the other.

377. Stationary Waves.—Lippmann's Colour-Photography.—Suppose a beam of parallel rays, or, in other words, a series of plane waves, is incident normally on a plane mirror, then the waves will be reflected at the mirror, and we may, as in the case of water waves (§ 275), have stationary waves set up owing to the interference between the direct and reflected waves. Consider a point P (Fig. 359), at a distance x from the mirror, then we may consider that at P we have two series of waves, one the direct waves and the other a series, which, starting in the same phase as the direct waves, has travelled along a path which exceeded that of the direct waves by $2x$. We must also add half a wave-length, for the reflection at O takes place at the surface of a denser medium. Hence the difference of path is really $2x+\frac{\lambda}{2}$.

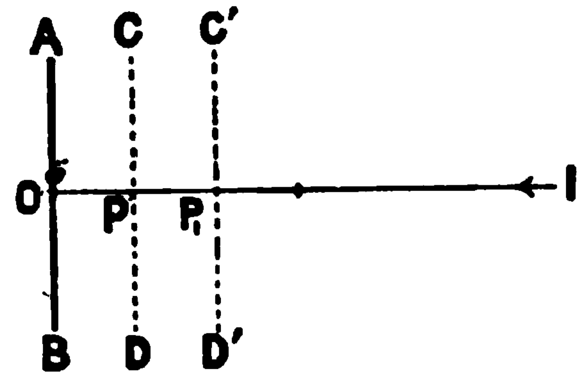


FIG. 359.

There will be interference at P, if this difference in path is equal to an odd number of half wave-lengths, or if

$$2x+\frac{\lambda}{2}=(2n+1)\frac{\lambda}{2}$$

$$x=\frac{n\lambda}{2}.$$

If then P is a point such that PO$=\lambda/2$, there will be interference throughout a plane CD, drawn through P parallel to the reflecting surface; there will also be interference throughout the plane C′D′, which is at a distance of λ from AB, and so on. The distance between the planes in which interference occurs will vary with the wave-length of the light, being smaller for violet light than for red light. The distance between consecutive planes, even for red light, is, in the case of normal incidence, excessively small, being only 3.8×10^{-6} cm. for the red (A line).

The formation of these planes, over which interference takes place, in the neighbourhood of a plane mirror has been utilised by Lippmann in his excessively beautiful method of obtaining photographs in natural colours. If the front surface of the mirror is coated with a sensitive photographic emulsion, then when light of wave-length λ' is incident on the mirror the light will be destroyed in planes which are at a distance of $\lambda'/2$, $2\lambda'/2$, $3\lambda'/2$, &c., from the mirror, so that the emulsion will not be affected on these planes. At the planes at distances $\lambda'/4$, $2\lambda'/4$, $3\lambda'/4$, &c., the incident and reflected light strengthen each other, and the emulsion will be acted upon, so that on development the silver of the emulsion will be reduced on these planes, and thus a number of parallel planes at a uniform distance apart will be produced within the film.

If now a beam of white light is incident on the developed plate, the

interval between each of the planes in which the silver has been deposited will act as a thin film producing interference in the manner considered in § 375 between the light reflected from two consecutive planes in which the silver is deposited. The difference in phase between the light reflected at plane 1 (Fig. 360), and that reflected at plane 2, will be equal to twice the distance between the planes. Hence, since the distance between the planes is $\lambda'/2$, the difference in the paths is λ', so that, in the case of the component of the incident white light which has the wave-length λ', the two reflected rays combine to strengthen each other. For light of all other wave-lengths the two reflected rays will differ in phase, and will therefore more or less interfere. If there were only two planes the selective strengthening of the reflected light of one wave-length would not be very marked; when, however, there are hundreds of planes placed one after the other, the final result is that practically only light of wave-length λ' is reflected, that is, light of the same colour as that originally incident on the sensitive film.

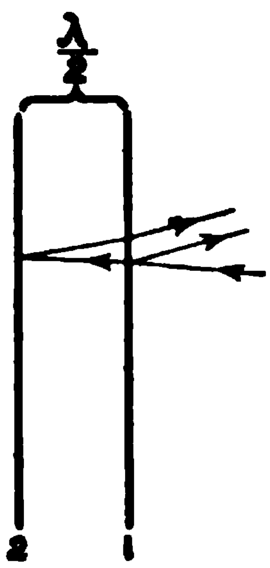


FIG. 360.

If then, instead of using homogeneous light to act on the sensitive film, light of different colours in different parts is used, such as would be obtained if the image of a party-coloured object formed by a lens is thrown on the film, then on development the silver will be so deposited that, when the film is afterwards illuminated by white light, the light reflected from different parts of the film will be of the same colour as the corresponding part of the original image, and hence of the object, and we shall thus get a photograph in natural colours.

378. Michelson's Interference Apparatus.—In the cases of interference which we have hitherto considered, the difference in the length of the paths of the interfering waves has only amounted at most to a few hundred wave-lengths. Michelson has, however, obtained interference when the paths differed by as much as 20 cm., *i.e.* about 400,000 times the wave-length. His apparatus consists of two parallel-sided plates of glass, G_1 and G_2 (Fig. 361), of equal thickness, and two plane mirrors, M_1 and M_2, arranged as shown in the figure. The surface of the glass plate G_1, which is turned towards the mirror M_1, is lightly silvered, so that when light is incident at an angle of 45° on this surface, half the light is reflected and half is transmitted through the thin coating of silver.

If a parallel beam of light is incident on the plate G_1 along the direction IO, the greater part will be refracted and traverse the plate. When this light meets the silvered surface half will be reflected, and after again traversing the glass plate will be incident normally on the mirror M_2; the other half will be transmitted, and after traversing the plate G_2 will be incident normally on the mirror M_1. The light which falls on the mirror M_1 will be reflected back along its path, will again traverse the

plate G_2, and will then be partly reflected at the silvered surface of G_1 along $O'R_1$. The light which falls on the mirror M_2 will be reflected back along its path and will traverse the glass plate G_1, and part will be transmitted through the silvering and emerge along the path $O'R_2$. In the figure the reflected rays are, for clearness, shown dotted and slightly displaced to one side of the incident rays. In reality no such displacement occurs, and the two reflected rays $O'R_1$ and $O'R_2$ are not separated.

The two rays start from the point O' in the same phase, and while one passes twice through the plate G_1 and twice traverses the distance between G_1 and M_2, the other passes twice through the plate G_2 and twice traverses the distance between G_1 and M_1. Hence, if the thickness of the glass plates is the same, the difference in phase of the rays $O'R_1$ and $O'R_2$ depends on the difference in the lengths of the paths $O'M_1$ and $O'M_2$. If

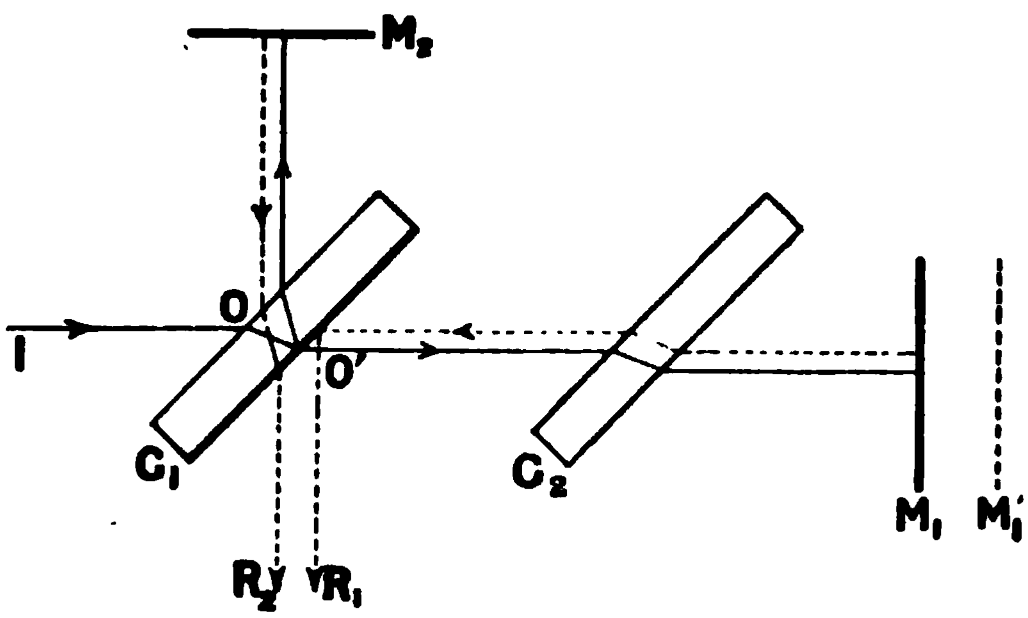


FIG. 361.

these paths differ by an odd number of half wave-lengths, the rays $O'R_1$ and $O'R_2$ will interfere.

Suppose that, for light of wave-length λ, interference takes place when the mirror M_1 is in the position shown, and that by means of a micrometer screw this mirror can be moved parallel to itself into a position M'_1, such that $\overline{M_1M'_1}$ is equal to half a wave-length, the path of one of the rays will be increased by a whole wave-length, so that if there was interference at M_1 there will also be interference at M_2, and nowhere between. Hence by moving M_1, and counting the number of times the two rays produce interference for any given wave-length, we shall be able to determine the distance through which we have moved M_1 in terms of the wave-length of the light used. Thus if we move M_1 through x centimetres, and interference is produced n times in this distance with light of wave-length λ, we shall have—

$$x=n\frac{\lambda}{2}.$$

Hence counting n, and knowing either λ or x, we can determine the other.

By means of this apparatus Michelson has compared the length of the *Metre des Archives* with the wave-length of light of certain colours. He used the three coloured lights given out by cadmium vapour, and found that if λ_R, λ_G, λ_B are the wave-lengths of the three cadmium lines in air under standard condition, then—

$$\begin{aligned} 1 \text{ metre} &= 1553163.6\lambda_R \\ \text{,,} \quad &= 1966249.7\lambda_G \\ \text{,,} \quad &= 2083372{\cdot}1\lambda_B, \end{aligned}$$

with a possible error of a few tenths of a wave-length. This measurement would allow us, supposing all the copies of the metre were destroyed, to reproduce the metre with a very high degree of accuracy.

379*. Explanation of the Rectilinear Propagation of Light on the Wave Theory.—One of the chief causes why the wave theory of light was for a long time thought to be incorrect, was the difficulty of explaining why light was propagated in straight lines, and did not, as sound in general does, spread out in all directions after passing through a hole in a screen; and we are now in a position to consider this question.

Let MM′ (Fig. 362) be the trace of a plane wave-front at right angles to the paper, and P a point at which we require to calculate the effect which will be produced by the wave. Now we may consider that each of the ether particles in the wave-front MM′ becomes a centre of disturbance, and we then have to find what is the combined effect of all these centres on the ether at the point P. From P draw PA perpendicular to the wave-front, and let the distance $\overline{PA}$ be called d. Next, with radii equal to $d+\lambda/2$, $d+2\lambda/2$, $d+3\lambda/2$, &c., describe a series of circles with P as centre, cutting MM′ at B, C, D, &c., and join BP, CP, DP, &c. Now since $\overline{BP}-\overline{AP}=\lambda/2$, the waves sent by the ether particles at A and B will reach P in opposite phases, and will therefore interfere. In the same way the waves sent from B and C will interfere, and so also the waves coming from the particles between A and B will interfere with the waves coming from the particles between B and C. Now the same argument will apply to the waves coming from all the particles on the wave-front included in a circle described about A as centre, and with radius $\overline{AB}$ on the one hand, and the particles included in the annulus or zone having radii $\overline{AC}$ and $\overline{AB}$. Now the effect produced at P by the waves sent from all the particles in any zone will depend on two things, namely, the area of the zone, which gives the number of ether particles which are sending waves to P, and the inclination of the

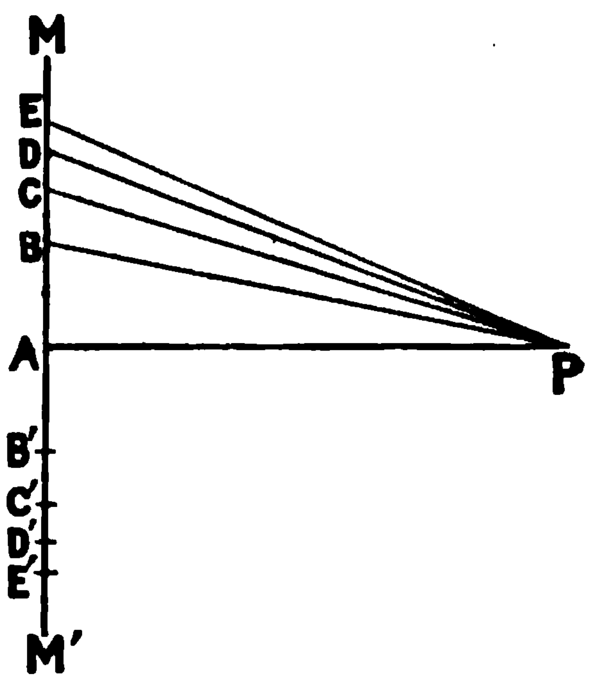


FIG. 362.

line joining P to the zone to the wave-front MM′. Since this inclination increases as the zones are taken further and further from A, the magnitude of the effect produced at P by zones having equal areas will on this account gradually fall off. Since the inclination varies from one zone to the next at first quite rapidly, but, as we shall see later, this change very soon becomes excessively small, it follows that the difference between the effects produced by equal areas of consecutive zones is at first considerable, but soon becomes inappreciable as we get away from A.

We have next to calculate the areas of the successive zones. Now $\overline{PB}=d+\lambda/2$. Hence

$$\begin{aligned}\overline{AB}^2&=(d+\lambda/2)^2-d^2\\&=d^2+d\lambda+\lambda^2/4-d^2\\&=d\lambda,\end{aligned}$$

if we neglect the term involving λ^2, since λ is a very small quantity. Also

$$\begin{aligned}\overline{AC}^2&=(d+\lambda)^2-d^2\\&=2d\lambda,\end{aligned}$$

and

$$\overline{AD}^2=3d\lambda,$$

and so on.

Hence the area of the circle AB is

$$\pi d\lambda.$$

The area of the zone BC is

$$2\pi d\lambda-\pi d\lambda=\pi d\lambda.$$

The area of the zone CD is

$$3\pi d\lambda-2\pi d\lambda=\pi d\lambda,$$

and so on. Hence the area of all the zones is the same.

Now if the distance $\overline{AP}$ or d is 10 cm., and λ is 5×10^{-5} cm. (green light), the radii of the zones have the following values :—

Zone.	Radius.	Width of Zone.
	Centimetres.	Centimetres.
1 ($\overline{AB}$)	0.022	0.022
2 ($\overline{AC}$)	0.032	0.010
3 ($\overline{AD}$)	0.039	0.007
4 ($\overline{AE}$)	0.045	0.006
5	0.050	0.005
...	...	...
10	0.071	0.003
11	0.074	...
...	...	...
100	0.224	...
101	0.225	0.001

This table shows very clearly how the width of the zones diminishes very quickly at first, and then more slowly, and how very narrow the

zones become even at a distance of two millimetres from the point A, which is called the pole of P.

Now the effect produced at P by any given zone depends on the area of the zone and on the inclination to the line AP of the line joining P to the zone. The table given above shows that for the zones at quite a short distance from the pole A the width of the zones is very small, and hence the angles between the lines joining two adjacent zones to P and the line AP are practically the same. Thus, except in the case of the first few zones, the effects of consecutive zones at P are exactly equal and opposite, and hence the only portion of the wave MM′, which contributes to the production of the disturbance at the point P, is that portion immediately surrounding the pole A. Thus if an opaque obstacle be placed at A, so as to cut off the disturbance coming from, say, the first ten zones, there will be no disturbance produced at P, for the disturbance coming from the zones, into which the rest of the wave can be divided, will neutralise each other by interference. That is, an obstacle of about 1.5 mm. diameter, if placed at A, will completely screen P. This result, of course, amounts to the same thing as the rectilinear propagation of light, for the obstacles employed when considering this phenomenon are in general larger than that given above.

We also at once see why it is that in the case of sound-waves "shadows" are so seldom formed. Thus, taking the case of a tuning-fork giving the note C of 512 vibrations per second, the wave-length in air is about 66.7 cm. Hence if the point P is at a distance of 1000 cm. from the pole A, the diameter of the tenth zone is $2\sqrt{10 \times 66.7 \times 1000} =$ 16340 cm. In other words, the diameter of an obstacle to shut off the sound would have to be more than sixteen times the distance of P from the pole, and, under these circumstances, the obliquity of the disturbance coming from the zones would be so great as to make our investigation only a very rough approximation. Thus we see that the reason we do not obtain sound-shadows is that the wave-length of the disturbance is too great compared to the size of the obstacles ordinarily used. Where the obstacle happens to be very large, sound-shadows are sometimes observed; as, for instance, an intervening hill has often protected certain buildings from the aerial concussion produced by an explosion, while other buildings at much greater distances, but not in shadow, have had their windows broken.

380*. Diffraction.—To complete the discussion of the production of shadows on the wave theory, we must now briefly consider the phenomenon observed in the immediate neighbourhood of the edge of a shadow, and also what happens when the size of the obstacle is less than the diameter of, say, ten zones, so that the disturbance is not completely cut off from the point P by the intervention of the obstacle.

We will first consider the case of a parallel beam of light which is intercepted by an opaque object, of which one edge is a straight line.

Let N (Fig. 363) be the section of the edge of the obstacle, NM′, taken at right angles to the paper, and P the point where the illumination is to be calculated. If no obstacle were present, we might divide the incident wave MM′ into half wave-length zones, just as in the previous section. Let the amplitude of the vibration which reaches P when no obstacle is present be A, so that the intensity of the illumination at P is A^2 (§§ 309, 359). Now, when the obstacle is so placed as to exactly cover half the zones, that is, when the edge passes through the pole of P, the amplitude of the disturbance produced at P will be reduced to a half, and therefore the intensity of the illumination will be $A^2/4$, that is, reduced to a quarter.

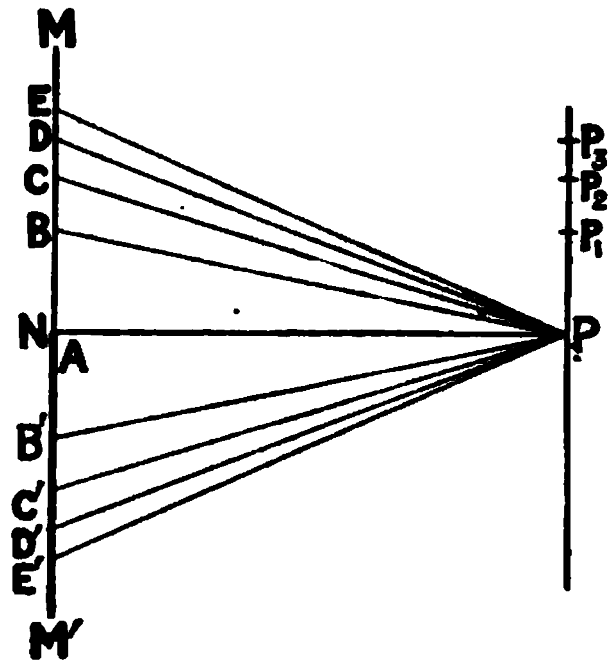


FIG. 363.

Now let the obstacle be gradually moved down till the edge N coincides with B′, that is, till the first zone is completely uncovered. The result will be that the illumination at P will increase and become considerably greater than A^2. The reason is that the illumination is A^2 when the zone B′C′ is also uncovered, and this zone affects P in the opposite phase, and therefore decreases the disturbance produced by the central zone. If the edge is now lowered to C′ the intensity of the illumination will gradually decrease, and reach a minimum value which is less than A^2, for the next most important zone, namely C′D′, is covered, and this would increase the disturbance at P if it were in action. Proceeding in this way, we see that the illumination at P will pass through a number of maxima and minima. The variation from the value A^2 will, however, become less and less as more zones are uncovered, and when about ten zones are uncovered, the illumination will remain constant at the value A^2.

When the edge has uncovered the first zone, the illumination at P will be the same as that at a point P_1, where $PP_1 = AB$, before the edge was moved from the position shown in the figure. Hence, since the illumination at P when the edge is at B′ is greater than A^2, it follows that the illumination at P_1, when the edge is at the point A, must also be greater than A^2. Thus if a screen be placed at P_3P, there will be a series of maxima and minima of illumination near the points P_1, P_2, P_3, &c., when the edge is at A.

Next, to examine the illumination which will be produced on the portion of the screen below P, that is, within the geometrical shadow of the obstacle M′N. When the edge is at A, the illumination at P is $A^2/4$, and as the edge is moved up the central zones are gradually covered, and hence the intensity of the disturbance sent to P gradually falls off.

The decrease in the illumination within the geometrical shadow is continuous, that is, there are no maxima and minima. The reason is that, supposing the edge to occupy the position shown in Fig. 364, then, starting from N, we may divide the remainder of the wave-front into half-period zones NB, BC, CD, &c. Of these zones, each will produce a greater effect than the next, but adjacent ones will send to P waves in opposite phase. Thus the illumination sent to P will be practically the difference of the effects of the first two zones, or at any rate of the first three or four. As the distance NA is increased, that is, as P is taken further and further inside the geometrical shadow, the difference between the effect produced by the first two zones will decrease, just as in § 379 we found that consecutive zones, after about the tenth *from the pole*, had equal and opposite effects.

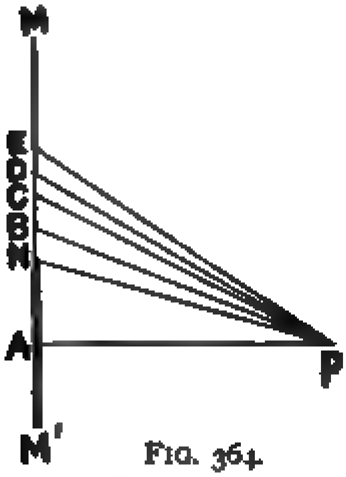


FIG. 364.

Thus the wave theory indicates that the shadow cast by a sharp edge when illuminated by parallel light, or, what comes to the same thing, light from a point, source, or narrow slit at a considerable distance, is not quite sharp. Outside the geometrical shadow will be a series of light and dark bands, and inside the light will not cease suddenly, but will fall off rapidly.

The intensity of the illumination on a screen placed at a distance of one metre from a diffracting edge, and illuminated by a parallel beam of light, is shown by means of a curve in Fig. 365. It will be seen that the

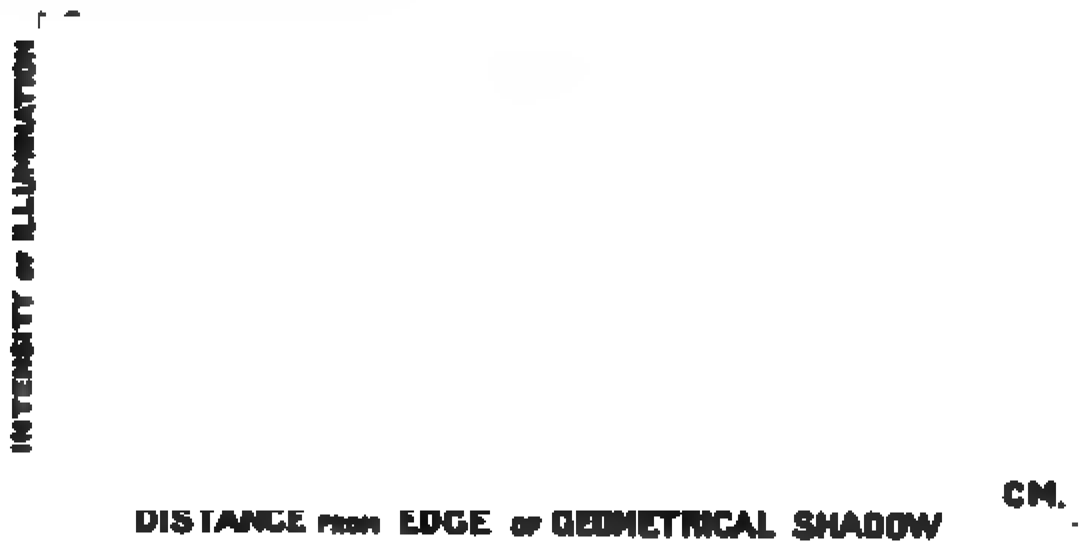


FIG. 365

illumination at the edge of the geometrical shadow is a quarter of the illumination at some distance from the edge, that is, of the illumination which would occur if the diffracting obstacle were removed.

If we have parallel light falling on a slit, then, as before, we may divide the incident wave into half wave-length zones with reference to a point P. If the slit is at a considerable distance from the pole of P, it will include many zones, for at this distance from the pole the zones are very narrow, and the total effect of these zones will be zero. As the slit is moved nearer to the pole, the number of zones included in the portion of the wave which can pass through the slit decreases, and when there is an even number the zones very nearly neutralise each other's effect, and there is a minimum of illumination at P, while when the slit includes an odd number of zones, the illumination is a maximum. The illumination will, of course, be a maximum at a point immediately opposite the slit. On either side will be formed a number of alternate dark and bright lines, the intensity of the maxima rapidly decreasing as we go away from the central band.

CHAPTER VIII

EMISSION AND ABSORPTION OF LIGHT

381. Nature of the Light emitted by a Luminous Body—Spectra.—In § 368 we have referred to the spectrum obtained when sunlight is passed through a prism, we now have to examine the constitution of the light given out by other sources.

If a solid body, such as a piece of lime or of metal, is heated, it begins to glow with a dull red colour at a temperature of about 600° C., and if the light emitted is examined in a spectroscope only the red end of the spectrum will be seen. At a temperature of about 1000° the yellow will appear as well as the red, while at about 1600°, the solid will glow with a white light, and the spectrum will stretch from the red to the violet.

The spectrum thus obtained with a glowing solid will differ from the solar spectrum in that there will be no dark bands, the spectrum being continuous from one end to the other.

The same character of spectrum is given by incandescent fluids, such as molten platinum.

When, however, the light given out by glowing gases or vapours is examined, the spectrum produced is of an entirely different character.

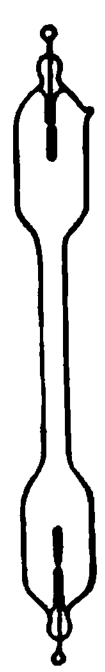


FIG. 366.

Thus, if a salt of either of the metals sodium, calcium, strontium, lithium, &c., is held in a colourless flame, such as that of a Bunsen burner, and the light is examined in a spectroscope, the spectrum will be found to be no longer continuous, but to consist of a number of bright lines in various parts of the spectrum. The position and number of these lines varies for the different metals, but does not depend either on the salt of the metal used (chloride, bromide, sulphate, &c.) or on the nature of the flame into which the salt is introduced. The *number* of lines visible with any given metal depends, to a certain extent, on the temperature of the flame, but although new lines may make their appearance as the temperature is raised, the position of the lines already present does not vary. In the case of gases, the spectrum is obtained by passing the spark from an induction coil (§ 524) through the gas which is contained in a rarefied condition in a tube of the shape shown in Fig. 366. In addition to line spectra, under certain conditions of pressure and temperature, the spectra of some gases exhibit bands of light,

which with a small dispersion are generally sharply defined on one side, but shade off gradually on the other. With a high dispersion, these bands are seen to be composed of numerous lines packed close together. When, however, the temperature is raised, the band spectrum becomes changed into a line spectrum.

The character of the lines in the spectrum of a gas depends very much on the pressure to which the gas is subjected. Thus in the case of hydrogen, at low pressures, say below 1 mm. of mercury, the spectrum consists of three narrow lines, one in the violet, one in the blue, and one in the red, which are generally indicated by H_γ, H_β, and H_α. As the pressure is increased, first the line H_γ, then H_β, and finally also H_α becomes wider, while under a pressure of about 36 cm. of mercury the spectrum is practically continuous. The explanation of these changes, if we accept the kinetic theory of gases, is comparatively easy. When a gas is under a low pressure, the mean free path (§ 141) of the molecules is great, so that the interval between successive impacts of a molecule with another is comparatively great. Thus although during the impact the atoms will be set into all kinds of forced vibrations, yet all these vibrations, except those which correspond to the *natural* period of vibration of the atoms, will very rapidly die out, and for the greater part of the time the atoms will be vibrating in their own natural period. Hence, if we suppose that in a glowing gas the light emitted is due to the vibrations of the atoms, it is evident that at low pressures the gas will give out light of certain definite wavelengths, corresponding to the natural periods of the atoms. As the pressure increases the mean free path of the molecules decreases, and hence the impacts become more frequent. Under these circumstances the forced vibrations will begin to tell, and at first it will be those vibrations which are nearly of the same period as the natural period that will be most noticeable, so that the bands will widen out. When the pressure is further increased, the encounters between the molecules are so frequent that the forced vibrations persist from one encounter to the next, and hence vibrations of all periods will be taking place in the different molecules, and a continuous spectrum will be obtained.

382. Series of Spectral Lines.—If we assume that the frequency of the light vibrations given out by a luminous body is the same as the frequency of the vibrations set up within the molecules of the substance, we are led to the conclusion that the motion of even a gaseous molecule must be very complicated, for the spectrum of most substances contains quite a large number of bright lines, each line corresponding, on the above hypothesis, to a different mode of vibration.

Although at first sight the arrangement of the lines in the spectrum of a gas or vapour appears in general quite irregular, yet a study of this subject has shown that in many cases certain relations are found to hold between the frequencies of the various lines.

The first relation of this kind observed is due to Balmer, who noticed that the wave-lengths, λ, of the lines in the hydrogen spectrum can be represented with great accuracy by the general expression

$$\lambda = 3645 \frac{m^2}{m^2 - 4} \times 10^{-6} \text{ cm.},$$

in which m is in succession given the values 3, 4, 5, &c., up to 16. The kind of agreement obtained between the observed values and those calculated from Balmer's formula is shown in the following table :—

m	Wave-length.	
	Calculated.	Observed.
3	6561×10^{-6} cm.	6560.7×10^{-6} cm.
4	4860	4859.8
5	4339.3	4340.1
6	4100.6	4101.2
7	3969	3968.1
8	3888	3887.5
9	3834.4	3834.0
10	3796.9	3795.0
11	3769.6	3767.5

Another curious fact is that when there exists in the spectrum of an element a doublet or triplet, that is, two or three lines close together, there are also, in general, a number of other doublets or triplets, and the difference between the frequencies of the components of these doublets and triplets is the same for all those which occur in the spectrum of any one element. Thus, in the case of thallium, Kayser and Runge have found the following values for the reciprocals of the wave-lengths of the components of the doublets. The reciprocal of the wave-length being proportional to the frequency of the vibrations, the differences will also be proportional to the differences of the frequencies.

$1/\lambda$	Difference.	$1/\lambda$	Difference.
18684.2 26476.6	7792.4	35372.1 43164.7	7792.6
28324.1 36117.1	7793.0	36879.2 44671.0	7791.8
30952.1 38744.8	7792.7	37503.0 45293.8	7790.8
33569.4 41365.1	7795.7	38305.0 46096.8	7791.8
34217.7 42010.2	7792.5	38663.3 46452.4	7789.1
34526.2 42321.4	7795.2	39157.0 46947.3	7790.3

If we plot the values of $1/\lambda$ for the lines given in the above table as abscissæ, as shown in the upper line of Fig. 367, where, since the components of the doublets would be at a constant distance apart throughout, only one has been plotted, the lines do not appear regularly arranged. If, however, the fourth and sixth lines are omitted, the remaining lines can be arranged in two series as shown at B and C, each of which resembles the series of lines represented by Balmer's formula, and can be represented by a similar formula. The separation of the

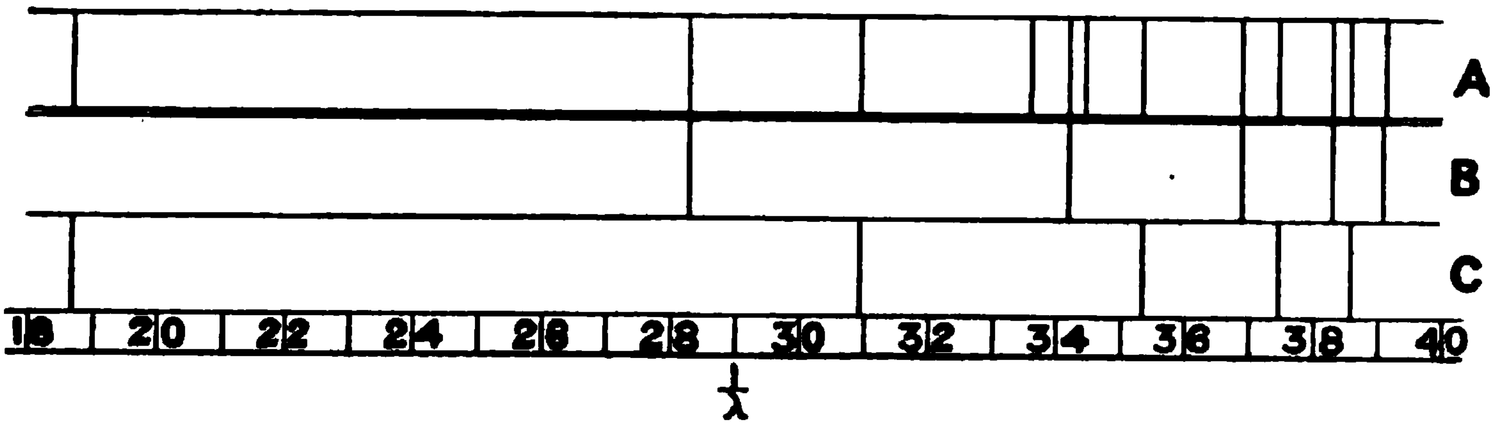


FIG. 367.

lines into two series is further justified by the fact that each double line of the first series is accompanied, on its more refrangible side, by a strong line which is easily reversed, while the lines of the second series are not accompanied in this way. The further consideration of this subject is beyond the scope of this work, but sufficient has been said to indicate the direction in which modern work on the classification of the lines in the spectra of the different elements is proceeding.

383. Absorption of Light.—When light passes through a medium, this medium in general absorbs part of the radiation, and the amount of this absorption is generally different for light of different wave-lengths, or, in other words, most media exert a selective absorption on light.

In order to examine the character of the absorption, white light is passed through the given substance, and the transmitted light is examined spectroscopically. If then the substance absorbs light of any particular wave-lengths more strongly than it does light of other wave-lengths, the spectrum will be crossed by dark bands corresponding to the colours which have been absorbed. Thus if a dilute solution of permanganate of potash is used, the spectrum is crossed by five dark bands in the green, while a dilute solution of human blood produces well-marked absorption bands in the yellow and green.

In the case of solutions, the absorption bands are generally fairly wide, the width increasing with the strength of the solution. When light is absorbed by gases or vapours, however, the character of the absorption bands is very different, the bands are sharply defined, and in general consist of a number of fine narrow lines in various parts of the spectrum. Thus if white light from a very hot body, such as the electric

arc, is passed through a flame which is strongly coloured yellow by means of sodium, two fine absorption bands are formed in the orange-yellow. If the white light is then cut off, the light from the sodium flame will give two bright lines, which occupy exactly the same place as did the dark absorption lines. We thus see that the light absorbed by the sodium vapour is of exactly the same wave-length as that which it itself gives out. This is really a case of resonance, for the incident white light contains waves of all periods; of these, the sodium molecules will most powerfully absorb those which are of the same period as their own natural periods. We may illustrate this action by taking the case of a number of ships at anchor, when if the period of the waves happens to coincide with the natural rolling period of the ships, then they will be set into violent oscillation, and the energy to set them into this oscillation having been derived from the waves, the waves must themselves have been absorbed, parting with their energy. Waves of other periods will, however, not set the ships into such violent oscillation, and hence will not be so strongly absorbed. In the case of the absorption by the sodium vapour, the white light, coming as it does from a source at a very high temperature, is very bright, and the sodium flame being at a much lower temperature, the sodium vapour absorbs the light of the wave-length it itself emits. Thus after traversing the flame all the constituents, except the yellow sodium light, exist in their original brilliancy; the sodium light, however, is only that due to the feeble radiation of the flame, so that by comparison with the light of the other colours the yellow sodium band looks black, although when the white light is cut off, so that there is no contrast, the sodium line is seen to be really bright.

When white light is transmitted through an incandescent gas, we therefore get dark bands in those parts of the spectrum corresponding to the bright lines produced by the light given out by the gas, this being a particular case of the general proposition that bodies absorb most strongly that kind of vibratory motion which they are themselves capable of giving out, whether it be water-waves as in the case of a ship, sound-waves as in the case of a resonator, or light and heat waves as in the cases just considered (Stokes' Law).

384. Reversal of Lines in the Solar Spectrum. — We have referred in § 368 to the black Fraunhofer lines in the solar spectrum, and from what has been said as to the reversal of the spectral lines produced by passing white light through a glowing gas, we are at once led to the explanation of these lines. They are due, as was first pointed out by Kirchhoff, to the absorption of certain portions of the light given out by the white hot nucleus of the sun during its passage through the gases and vapours which surround this nucleus, or through the earth's atmosphere.

Since the position of any one of Fraunhofer's lines in the spectrum

coincides with the position of a bright line, due to the vapour which has absorbed the light, by measuring the wave-lengths of Fraunhofer's lines and comparing them with the wave-lengths of the bright lines produced by the elements which occur on the earth, we can discover whether these elements occur in the sun's atmosphere.

Thus the Fraunhofer line D in the solar spectrum really consists of two lines close together (D_1 and D_2), and these occupy the same positions in the spectrum as two of the lines of the metal sodium, so that we may infer that sodium exists in the vaporous condition in the sun's atmosphere. Until recently certain lines in the solar spectrum were unknown amongst terrestrial elements, and were said to be due to an unknown element, helium. This element, which proves to be a gas, has, however, now been discovered, so that in this case we may almost say that this element was recognised on the sun before it was known on the earth.

385. Displacement of Spectral Lines.—We have seen in § 296 that when a sounding body is either approaching or receding from an observer, the pitch of the note perceived is different from that given out by the sounding body. The same principle applies in the case of light. Thus if a luminous body, in the spectrum of which there are definite lines, is moving towards the observer, the wave-length of each of the lines will be apparently shortened, since in a given time the observer will receive more waves than he would if the luminous body were stationary. The effect of this is that the lines in the spectrum will be displaced towards the violet end of the spectrum. In the same way, if the source of light is moving away from the observer, the lines will be displaced towards the red end of the spectrum. Hence, by comparing the position of the lines in the spectrum of a star with the position of the same lines in the solar spectrum, we can determine whether the distance between the earth and the star is decreasing or increasing, and from the extent of the displacement of the lines we can calculate the velocity with which the earth and the star are moving relatively to one another in the line joining the two. In this way it has been found that Arcturus is approaching the earth with a velocity of 42 miles per second, while Aldebaran is receding with a velocity of 45 miles per second.

The same method has been used by Keeler to prove that Saturn's rings are composed of small bodies rotating round the planet. An image of the planet is formed on the slit of a spectroscope, so that the different parts of the spectrum, taken at right angles to the direction of the dispersion, correspond to light coming from the different parts of the planet along the line in which the slit cuts the image of the planet. If then the rings rotate, the spectral lines corresponding to the light from the two ends of a diameter of the rings will be displaced in opposite directions, while, if the outside of the ring rotates faster than the inside, the lines will not only be displaced as a whole, but, since the amount of the displacement depends on the velocity, they will be inclined. If the

outside of the rings is rotating more slowly than the inside, the lines will still be inclined, but in the opposite direction. By this method Keeler finds that the inside of the ring is rotating faster than the outside. If the ring were fluid or solid, the outside would move faster than the inside, while, if it consists of a swarm of independent solid bodies, the nearer ones will have to rotate the faster, or otherwise the centrifugal force would not be sufficient to keep them from falling into the planet, so that the spectroscope indicates that the latter hypothesis is the correct one.

386. Anomalous Dispersion.—We have seen, when speaking of dispersion (§ 369), that the dispersive power is not the same for all substances. For this reason the spectra produced by prisms of different materials are not similar, for the relative spreading of the different colours is not the same. Thus in Fig. 368 are given the relative positions of some of Fraunhofer's lines for spectra produced by prisms of different substances, the dispersion between the A and H lines being the same for all. The top line is a grating spectrum, which is added for the sake of comparison, for here the spectrum is normal, in that the dispersion

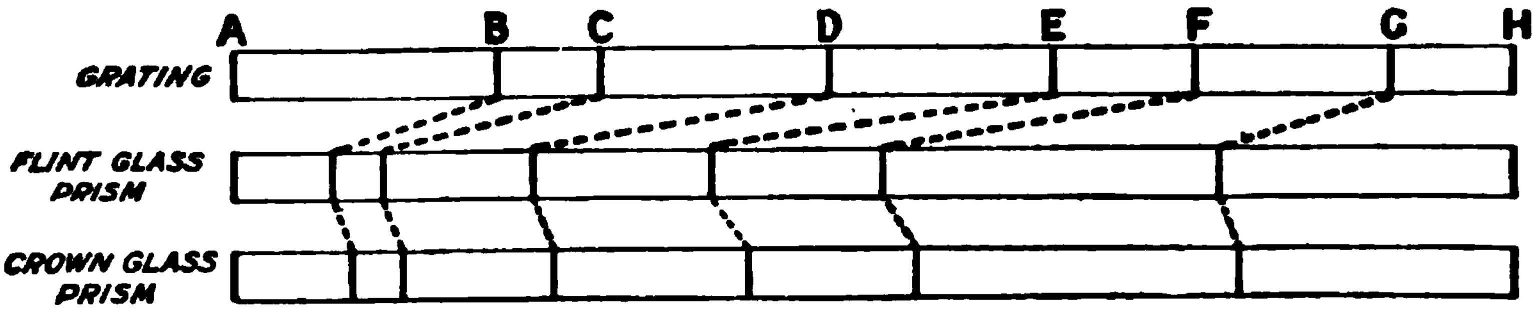


FIG. 368.

between any two rays is proportional to the difference between their wave-lengths, so that the dispersive power for all gratings is constant.

In the case where the material of which the prism is constructed shows marked selective absorption, the spectra obtained are very abnormal, for in certain cases the order of the colours is altered, so that these no longer follow in the order of their wave-length, while in other cases the spectrum, instead of being continuous, is separated into isolated parts by broad dark bands.

A solution of fuchsine (one of the aniline dyes) in alcohol strongly absorbs the green light, so that the spectrum formed by transmission through a prism of this substance does not contain any green. Of the three colours, red, orange, and yellow, on one side of the missing green, the red is least deviated, next the orange, and then the yellow; these colours following each other in the usual order. The deviation of the violet is, however, quite abnormal, for light of this colour is *less* deviated than the red, being separated fom this latter by a dark band.

In Fig. 369 the top line shows the arrangement of the colours, as indicated by the Fraunhofer lines, in the spectrum produced by a glass

prism, while the second line shows the arrangement of the colours in the spectrum produced by a prism filled with a solution of fuchsine. The last line is a curve, such that the ordinates represent the intensity of the various coloured lights in the fuchsine spectrum.

The light reflected from a solution of fuchsine at normal incidence is coloured green, thus accounting for the absence of the green in the transmitted spectrum.

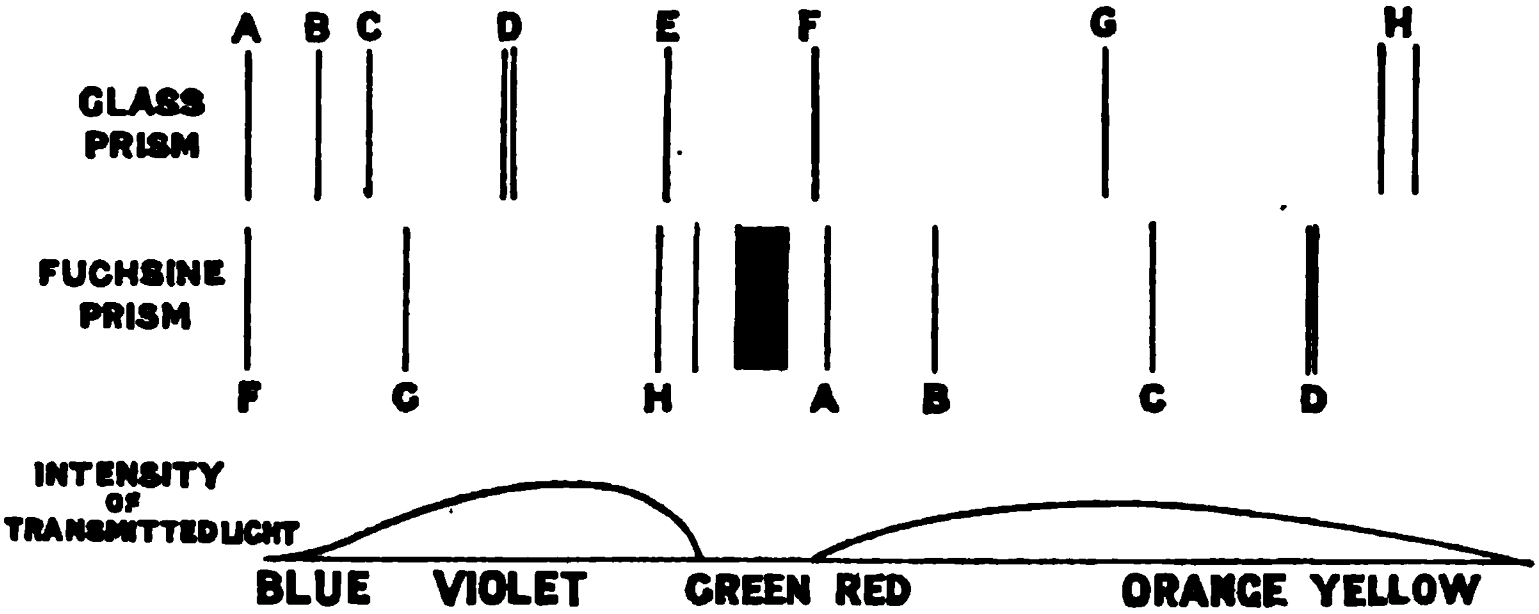


FIG. 369.

Kundt has made observations on a number of bodies which show anomalous dispersion, and he finds it in all bodies which have what is called surface colour, *i.e.* those whose colour, as seen by reflected light, is different from that seen by transmitted light. As a result of his experiments, he found that if we go up the spectrum in the sense of decreasing wave-lengths (*i.e.* from red to violet), the deviation is abnormally increased below an absorption band, and diminished above the absorption band. Thus in the case of fuchsine, which has an absorption band in the green, the colours red, orange, and yellow, which are below the band, are deviated to an abnormal extent; and the blue and violet, which are above the absorption band, are less deviated than the normal. This is shown graphically in Fig. 370. The dotted curve represents the proportion of light of the different wave-lengths which is absorbed and has a well-marked maximum in the green. The full-line curve ABCD represents the deviation produced by a given prism of fuchsine. The

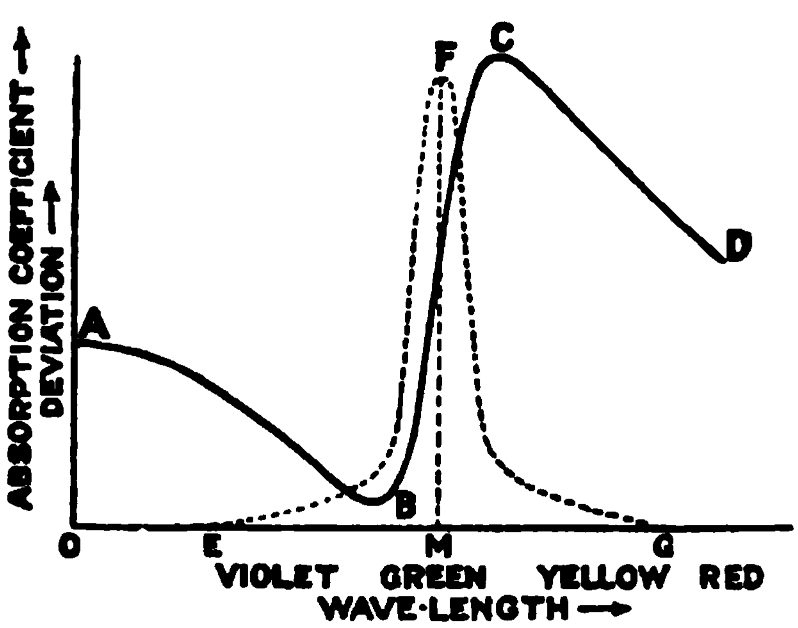


FIG. 370.

yellow and red, as shown by the portion DC of the curve, are deviated more than the violet, as shown by the portion BA.

387. Colour Produced by Absorption.—We have already referred to the fact that non-luminous bodies are seen by the light which they scatter at their surface. The light is usually not only reflected at the surface, but penetrates some depth below the surface, where it is reflected by inequalities in the substance of the body. This internally scattered light then reaches the eye, after traversing a certain thickness of the substance; and hence if the body absorbs any given coloured light, the reflected light will be of the tint obtained by removing light of this colour from white light. Thus a red poppy appears red because the petals exert a strong selective absorption on all the colours except red, so that light which has penetrated beneath the surface of the petals, and is then scattered by the cells, emerges robbed of all its colours except red. The fact that the flower absorbs almost completely all colours except red can be observed by holding the flower in different parts of the spectrum. In the red it will appear brilliantly red, and as it is moved towards the green the brilliancy will gradually fade, till, when it is in the blue, all the light which now falls on it will be absorbed, so that none is reflected to the eye, and the flower will appear black.

In the case of the bodies referred to in the previous section as showing surface colour, light of a particular colour seems unable to penetrate at all, and is therefore reflected, so that the transmitted light will be without this colour. Such phenomena are shown by many of the aniline colours, and by some metals, such as gold and copper. Thus in the case of gold the reflected light is of the colour ordinarily associated with the metal; a thin film of gold is, however, transparent, the transmitted light being green.

388. Distribution of Energy in the Spectrum.—A black body appears such because it absorbs light of all wave-lengths, and although even lamp-black reflects a little light, yet it absorbs such a great proportion of the incident light, that for most purposes it may be taken as absorbing the whole of the incident light. If, then, a beam of light is incident on a surface coated with lamp-black, the light will neither be reflected nor transmitted. The energy of the incident light-waves will be transferred to the absorbing body, where it will appear as heat, so that the body coated with lamp-black will become heated; and if we measure the quantity of heat it receives in a given time, when absorbing a given quantity of light, the energy corresponding to this incident light can be calculated. Hence by measuring the heat received by a black body when placed in different portions of a spectrum, the relative energy of the light of the various wave-lengths present in the spectrum can be determined, the instruments used for this measurement being the ones described in § 244.

In order to carry out this experiment, Langley uses a rock-salt prism

to produce the dispersion, since this substance is not only transparent to the light rays, but also to heat rays. In order to measure the energy in different portions, he uses a bolometer, the receiving surface of which consists of a very thin and narrow strip of blackened platinum.

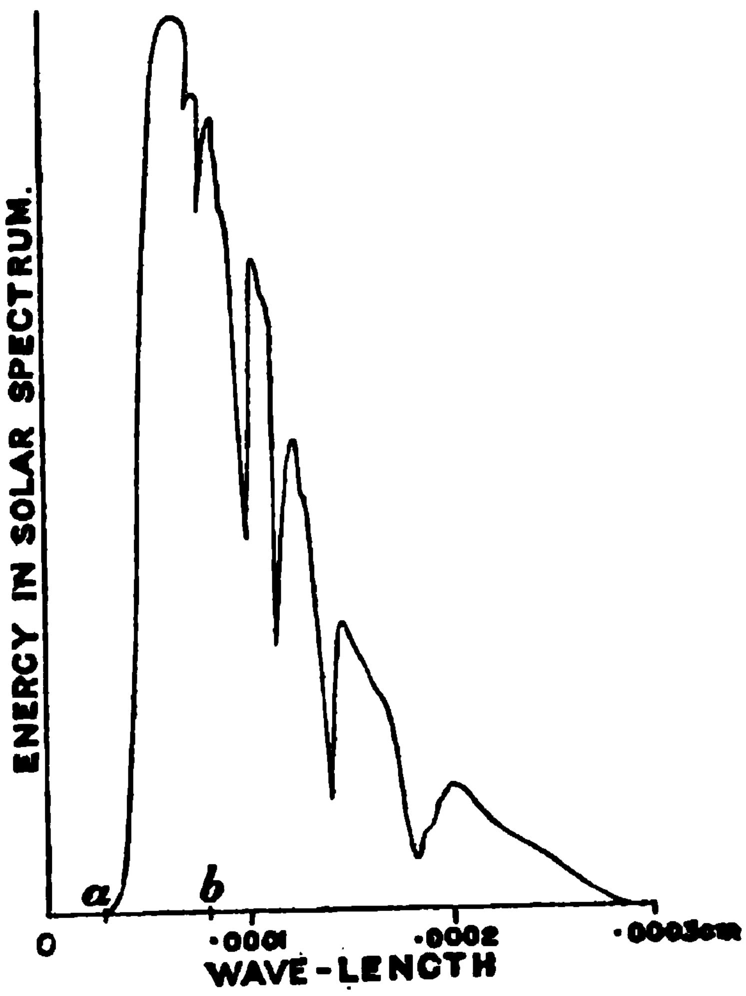


FIG. 371.

The results obtained in the case of the solar spectrum are shown in Fig. 371, in which the ordinates represent the energy of the light of the various wave-lengths. The part of the curve between *a* and *b* corresponds to the visible spectrum. It will be seen from this figure that the spectrum extends far below the extremity of the visible red.

Langley has also observed, by the same method, the radiation from bodies heated to temperatures below that at which they emit visible rays, and some of his results for blackened copper heated to various temperatures are shown in Fig. 372, together with the solar

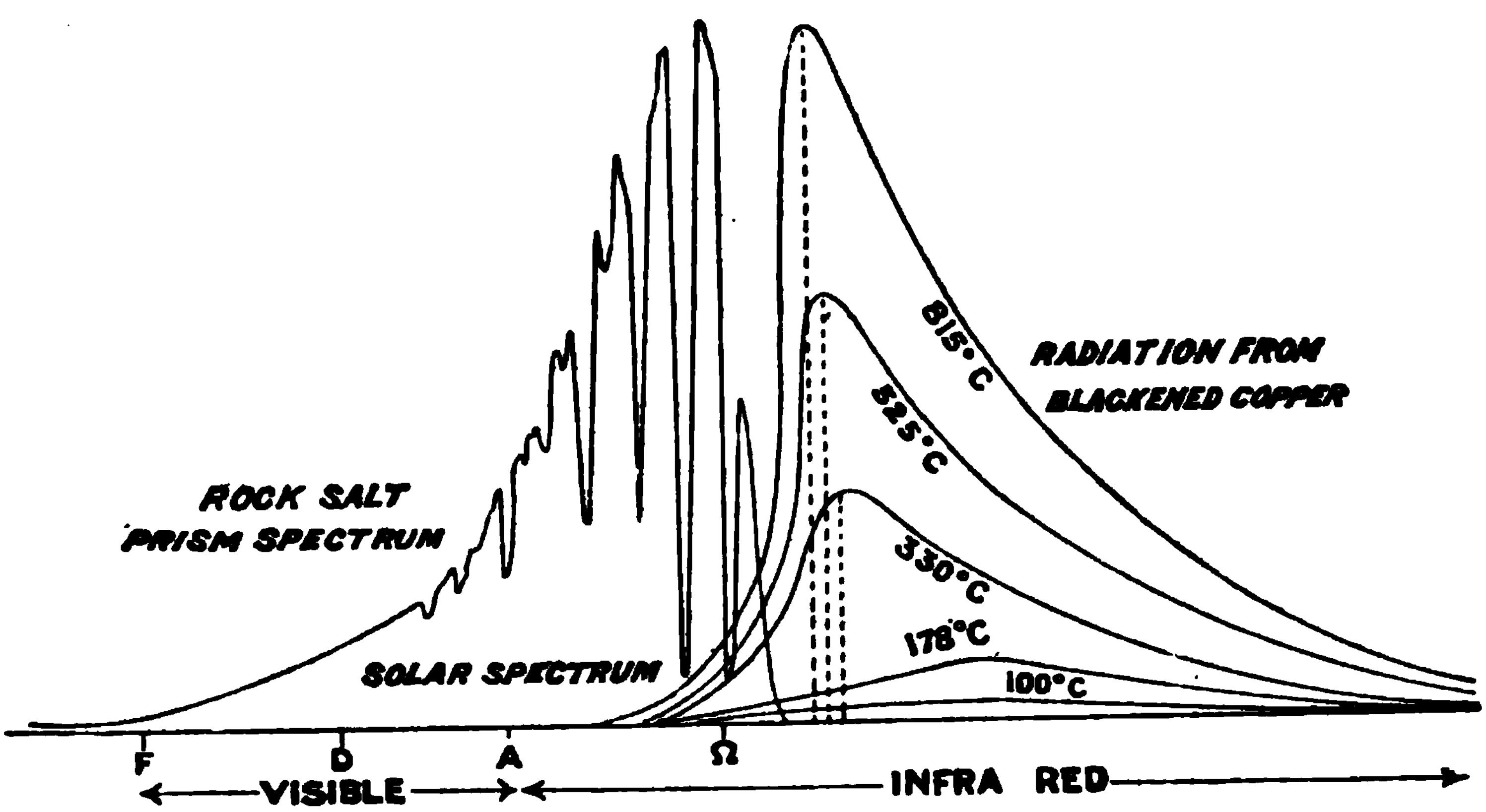


FIG. 372.

spectrum for comparison. In this figure the abscissæ represent the deviation produced by the rock-salt prism, and not wave-lengths, as in

Fig. 371. It will be observed that, as the temperature of the blackened copper is raised, radiation of smaller and smaller wave-length appears; and although the energy corresponding to each wave-length increases, the increase is not the same for all, the maximum of the curve being displaced towards the smaller wave-lengths.

Since when the copper is heated to a sufficiently high temperature it emits white light, we see that, starting at a low temperature, the radiation emitted is of comparatively great wave-length only, but as the temperature rises the body emits rays of smaller wave-length, till finally even violet light is emitted. We see, therefore, that there is no physical difference, except that of the wave-length, between the radiation given out by a hot kettle, which we call radiant heat, and that given out by a gas flame, which we call light. Hence the arbitrary distinction between heat-rays and light-rays, based on the fact that, although both are simply wave-motions in the ether, we perceive them with *different* senses is misleading.

The depressions in the energy curve of the solar spectrum in the infra red, indicating as they do the partial absence of radiation of certain wave-lengths, show that even in this region there are dark bands in the solar spectrum. These bands in the infra red are in a great measure due to absorption within the earth's atmosphere.

Not only does the solar spectrum extend far beyond the visible spectrum in the direction of increasing wave-lengths, but it also extends beyond the violet. We shall see later on how this ultra-violet portion of the spectrum may be examined.

389. Fluorescence.—If a solution of chlorophyll (the colouring-matter of green plants) is placed in a dark room, and a beam of white light is allowed to fall on it, the portions of the solution on which the light first falls become luminous, emitting in all directions a red light. This phenomenon is called fluorescence, the name being derived from fluor-spar, a body which also exhibits the phenomenon. The fluorescence is most brilliant at the surface of incidence of the white light, the brilliancy gradually decreasing with the thickness of the solution through which the light has passed. The same phenomenon is exhibited by paraffin oil, solutions of quinine, and of some aniline colours, such as eosin (red ink), fluorescine, also by some salts, such as barium, or potassium platino-cyanide.

If a fluorescent body, instead of being placed in white light, is exposed to light of different colours, it is found that the fluorescence only occurs with certain kinds of light. Thus if a test-tube containing a solution of sulphate of quinine is held in different parts of the spectrum, it presents a very different appearance in some parts from, say, the same test-tube when filled with water. In the red the solution of quinine looks red, in the yellow and green it looks yellow and green respectively, but in the blue and violet a marked change is apparent, as it begins to show the

pale blue fluorescent colour which it exhibits in white light. This fluorescence increases towards the violet end of the spectrum, and is visible even when the test-tube is held beyond the limits of the visible spectrum, so that the ultra-violet rays are capable of exciting fluorescence in this substance.

If a solution of chlorophyll is treated in the same way, it will be found to glow even in the red with a deep red. As the solution is moved up the spectrum, it continues to exhibit the *red* fluorescent light, although in the violet the fluorescent colour is brownish, due to the presence of some green light as well as the red.

If the light emitted by a fluorescent body is examined spectroscopically, it is found not to be monochromatic, but to contain light of various colours; the wave-length of these colours is, however, *always* less than the wave-length of the light which causes the fluorescence. Thus fluorescent bodies possess the property of absorbing light of certain wave-lengths; quinine absorbs most of the ultra-violet light, chlorophyll has a marked absorption band in the red, as well as others in the yellow, green, and blue, and of converting this absorbed light into light of greater wave-length. For instance, to the eye a beam of sunlight does not seem reduced in intensity by passage through a moderate thickness of a solution of quinine, but it has been deprived almost entirely of its ultra-violet rays, and the quinine has converted these rays into blue and violet rays which are visible to the eye.

Fluorescence has been used to map the solar spectrum beyond the violet, for when the spectrum is thrown on a fluorescent substance the fluorescent glow will appear extending beyond the violet and is traversed by dark absorption bands which are similar to the Fraunhofer bands in the visible part of the spectrum.

390. Phosphorescence.—In the case of the fluorescent bodies just considered, the emission of the fluorescent light ceases immediately the incident light is cut off. Some substances, particularly the sulphides of calcium, barium, and strontium, continue to emit light after the incident light has been cut off, so that after exposure to light they shine in the dark. This phenomenon is called *phosphorescence*, a name which is rather misleading, since the glow exhibited by phosphorus is due to slow chemical action, while the glow of a phosphorescent substance is not due to chemical action, and is really fluorescence which persists after the source of illumination is removed. Phosphorescence is exhibited by a great number of bodies, but in most cases it lasts for such a short time after the incident light has been cut off as to require special means for its detection.

391. Calorescence.—The converse action to that which occurs in the case of fluorescence, *i.e.* the absorption by a body of radiation of one wave-length and its emission as radiation of shorter wave-length, is called calorescence, and was exhibited by Tyndall by focussing the infra-

red rays from an electric arc, the luminous rays being removed by transmission through a solution of iodine, on a strip of platinum foil, when the platinum was heated to incandescence and emitted visible radiation.

392. Chemical Action.—When light is absorbed by a body the energy of the absorbed radiation is taken up by the body, and we have already considered some of the forms under which this absorbed energy can exist, namely, it can be converted into heat and warm the body, or it can produce by fluorescence or phosphorescence light of different character from the incident radiation, and be again radiated as a vibratory motion. We have now to consider a third way in which the energy of absorbed light may be used, namely, in doing chemical work. Thus a mixture of chlorine and hydrogen gases will keep indefinitely in the dark, but as soon as the mixture is exposed to sunlight the two gases combine with explosive violence.

Another well-known case where light produces chemical change is that of silver chloride, which, under the influence of light (sunlight), becomes blackened owing to the reduction of the silver. This provocation of chemical change by light is, of course, the cause of all photographic action.

Light of all colours is not equally active in promoting chemical change, and in Fig. 373 the relative intensity with which light of different colours

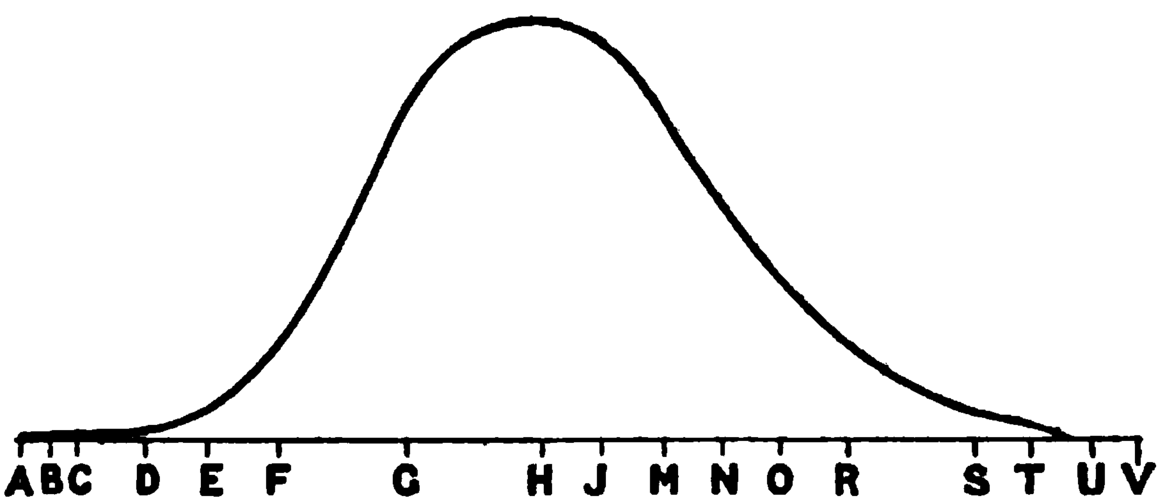


FIG. 373.

of the solar spectrum act in promoting chemical change is shown by means of a curve. The exact position of the maximum, however, depends to a certain extent on the nature of the chemical change produced. The spectrum must be produced by means of a quartz prism and lenses, for glass exerts a powerful absorption on the ultra-violet rays. It will be seen that the rays that are chiefly efficacious are those in the extreme violet and in the ultra-violet. For this reason the ultra-violet rays are often called chemical rays, but it must be remembered they only differ from the visible and heat rays in their wave-length, and that chemical action is not confined to these rays, but is only more strongly exhibited by them than by rays of greater wave-length.

393. Extent of the Light and Heat Spectrum.—It is of some interest to briefly collect a few data as to the range of wave-lengths which have been measured in the case of light or heat radiation.

In the following table the approximate wave-lengths of a few interesting kinds of radiation are given :—

	Cm.
Smallest wave-length measured	.00001
Maximum of chemical action in solar spectrum. Extremity of visible spectrum	.00004
D line	.000059
Extremity of visible spectrum	.00007
Maximum energy in solar spectrum	.00008
Largest wave-length measured	.0025
Smallest measured electrical oscillation	.6

The last number of the above table has been added on account of the fact that, on the electro-magnetic theory of light, both light and heat waves are really electrical oscillations (§ 581) of small wave-length, and the numbers given show that there is not such a very great breach to be filled up before we have a series of measured wave-lengths, *i.e.* a spectrum, extending continuously from the ultra-violet down to electrical oscillations which are observed as such.

CHAPTER IX

COLOUR SENSATIONS

394. Sensations produced by Light.—We have up to the present considered the subject of light in its objective aspect only, and must now proceed to examine the subjective sensations produced when light-waves of various kinds enter the eye. When dealing with the subject of audition, we saw that the nature of the sensation produced depends on the intensity, the frequency, and the timbre of the note. The timbre, however, is really included in the first two, for it depends on the frequency and intensity of the various simple tones which build up the note. In the same way, the sensation experienced when light enters the eye depends on the frequency and intensity of the various simple coloured lights which build up the resultant colour. As we shall, however, find, the eye possesses much less analysing power than the ear, for while the ear which receives a note of given composition can always distinguish any other note of which the composition is different, this is not the case with the eye. Thus by allowing light of only three selected frequencies to enter the eye, a sensation is produced which is quite undistinguishable from the sensation produced when white light enters the eye, although, as we have seen, white light consists of light of all frequencies between very considerable limits. Thus while in acoustics like sensations are produced by like causes, in optics this is not necessarily true, and the same sensation may be produced by entirely different causes.

395. Colour Constants.—We shall in the following pages use the word colour in a rather different sense to that hitherto employed. Up to now, by the colour of light we have meant the frequency of the ether vibrations, and so have used it in the sense of pitch in acoustics. Now, however, we shall use the word colour to designate the sensation produced in the eye, although where confusion is likely to occur the expression *colour sensation* will be employed.

As we shall see later, the colour sensation produced in different persons by the same quality of light may vary considerably, and so we have to consider the sensation which is felt by the majority of people; in other words, we shall deal with the normal eye.

In order to specify a colour it is necessary to know three things about it. In the first place, we require to know the frequency of the various vibratory motions which constitute the light which enters the eye, or, as

it is sometimes called, the hue of the light. In the second place, we require to know the brightness or luminosity of the colour. In the third place, we require to know whether the light considered is mixed with any white light, and if so, to what extent. If a light is free from admixed white light it is said to be pure. Thus by allowing monochromatic *D*-light to fall on a white card, the sensation is that of orange-yellow and the colour is pure. If, however, the card is simultaneously illuminated by white light, the sensation produced is altered and the colour is no longer pure.

396. Luminosity.—In order to be able to measure the luminosity of a colour, we must have a standard or unit of luminosity. Two cases have to be considered, namely, when we are dealing with the luminosities of different coloured lights, and when we are dealing with the luminosities of the colours seen when different pigments are illuminated by white light.

When dealing with lights of different colours, the unit taken is some of the white light produced by the source which is employed to give the coloured light. In the case of pigments, the unit is the luminosity of a white surface which is illuminated by the same light as that which falls on the pigment.

As a source of white light which may be employed in colour measurements, Captain Abney has found that the light given by the crater of the electric arc (§ 496) is by far the most steady and uniform in quality. The arrangement he has employed in his experiments on colour is shown in plan in Fig. 374. An image of the crater of an arc, E, is thrown, by means of a lens, L_1, on the slit of a collimator. The parallel beam of light thus produced falls on the prisms P_1 and P_2, and is thus split up, the different coloured rays being brought to a focus by the lens L_2 between V and R. If a screen is placed at A, a pure spectrum will be

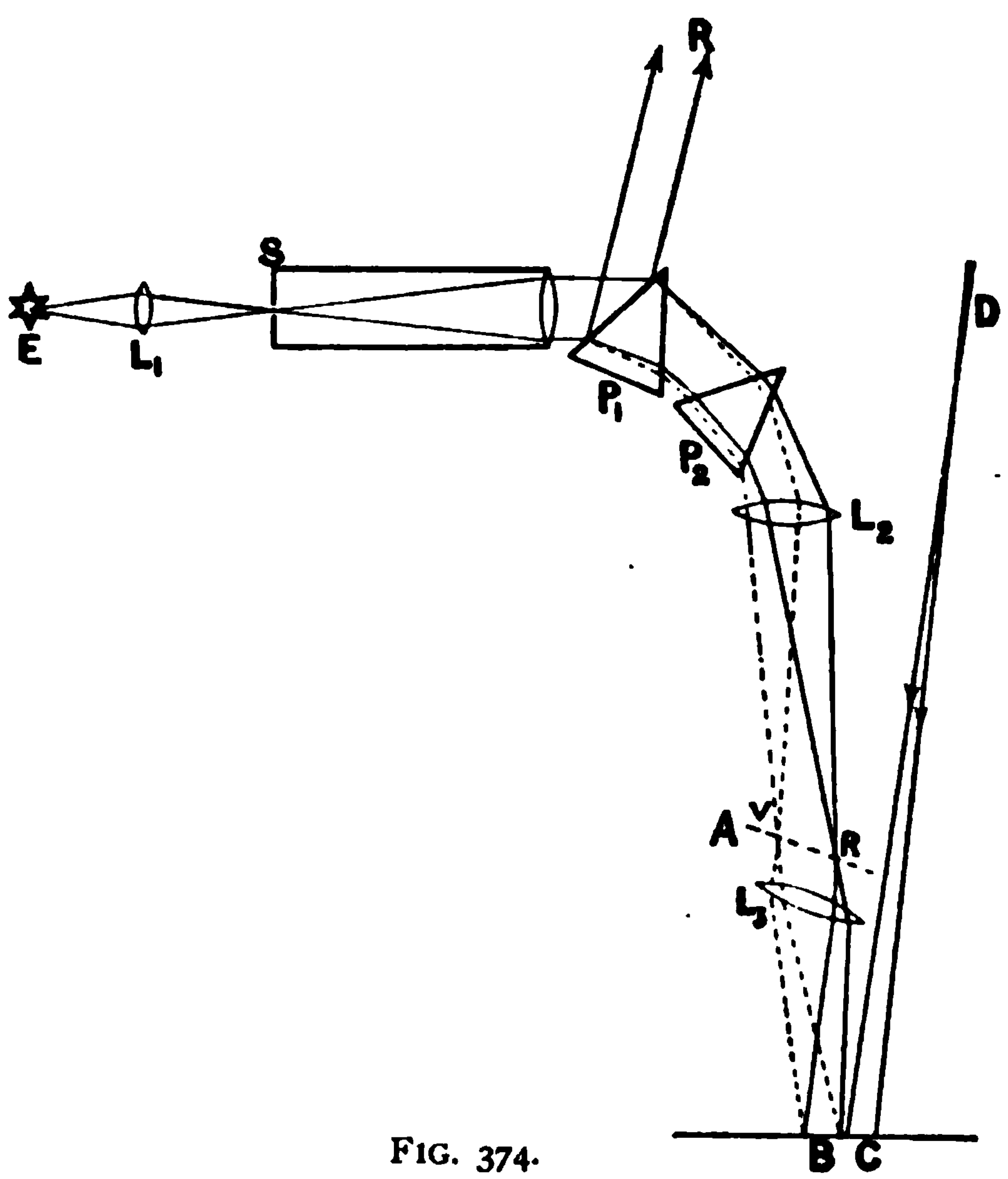


FIG. 374.

formed on it. If the screen is not there, the lens L_3 will cause the light of all the various colours to be superposed over a small patch on a screen at B, and so will reproduce white light. If, however, there is a screen at A in which there are one or more slits of which the positions can be varied, then it is only the light of the wave-lengths corresponding to the positions of these slits in the spectrum which will be thrown by the lens L_3 on the patch B. Hence by varying the positions and sizes of these slits, different colour mixtures can be obtained. Some of the white light is reflected from the first prism along R, and by means of a lens and a mirror this light is caused to form a white patch on the screen at C, and this acts as a reference white when measuring luminosities. The intensity of the white light can be reduced by means of a set of rotating sectors placed at D, which are so arranged that the proportion of opaque sector to transparent sector can be adjusted while the instrument is rotating.

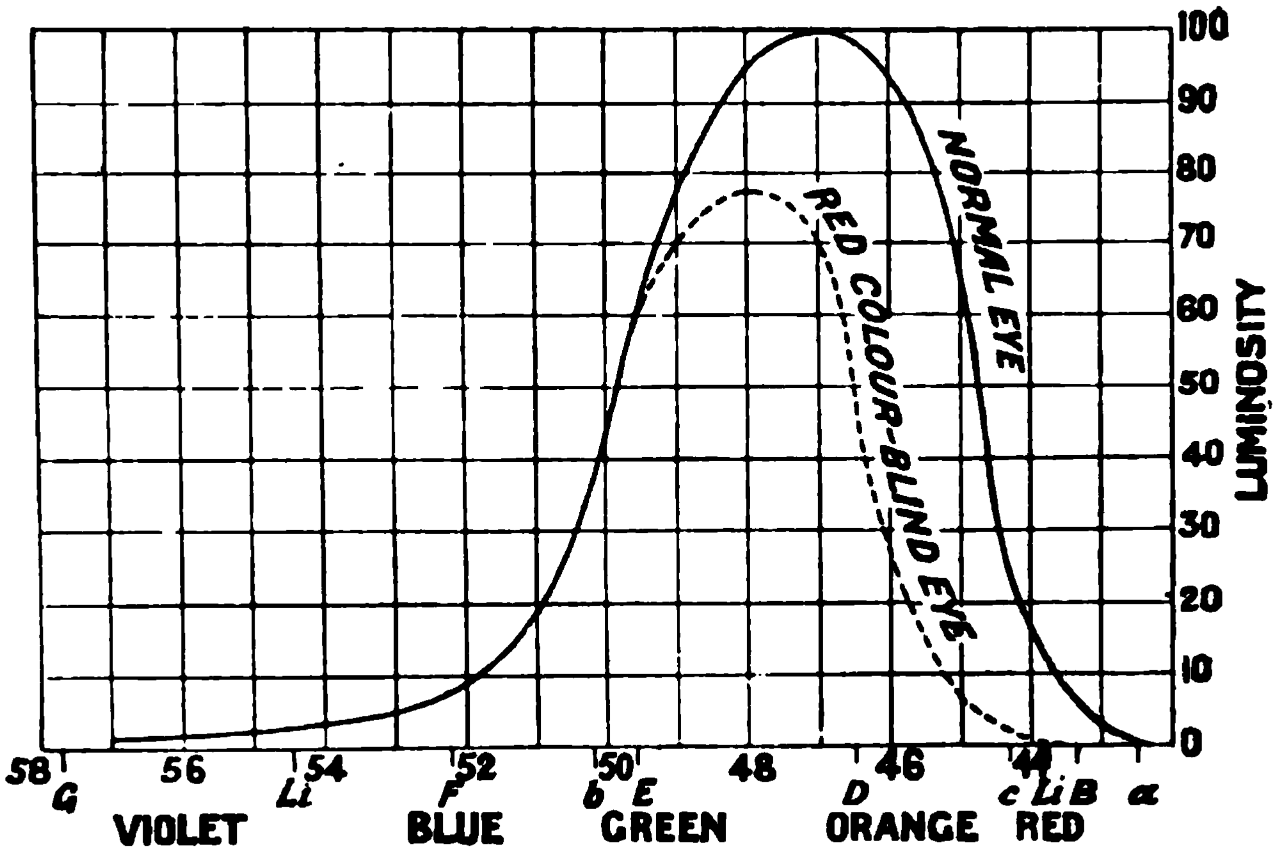


FIG. 375.

In order to determine the luminosity of the different parts of the arc-light spectrum a single slit is employed, so that the patch B is illuminated by light of one wave-length, and the intensity of the white light is varied till the two patches appear of equal brightness. By making this comparison all along the spectrum, and plotting the luminosities obtained as ordinates, Abney has obtained the curve shown in Fig. 375. The full-line curve represents the luminosity for a normal eye, and it will be seen that there is a very marked maximum in the yellow, and that the luminosity of the violet end of the spectrum is very small. The dotted curve represents the luminosity as measured by a red colour-blind observer, and we shall return to this subject later.

If the luminosity of two coloured lights is measured separately by

placing a slit at the appropriate place in the spectrum, and then two slits are placed so as to allow light of the two colours to fall on the screen simultaneously, and the luminosity of the two combined measured, it is found to be equal to the sum of the luminosities of the two separate components. The same result is obtained with three or more different coloured lights, so that the luminosity of such coloured lights is additive.

In order to measure the luminosity of the light reflected from various pigments, the arrangement shown in Fig. 376 is employed. The central portion, A, is covered with the pigment, and the proportion of the black and white sectors C and B can be varied by slipping one over the other. This card is mounted on an axle and rapidly rotated, and the width of the white sectors is altered till, when illuminated with any given coloured light, the luminosity of the whole disc appears the same. The ratio of the white sectors to the whole circle then gives the luminosity of the pigment when illuminated by the given coloured light.

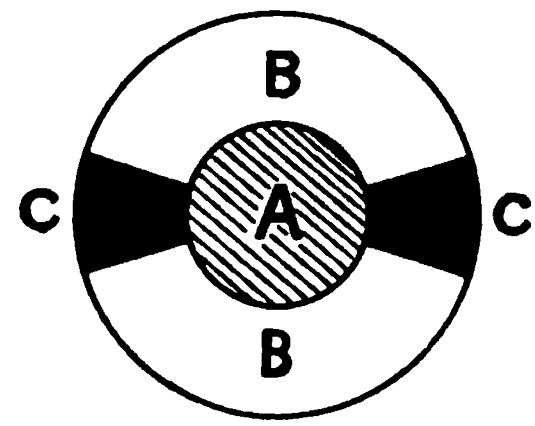


FIG. 376.

397. Colour Mixtures.—The apparatus shown in Fig. 374 can be used for studying the colour sensation produced by mixing light of various wave-lengths. For by placing slits in various parts of the spectrum the patch B will be illuminated by the mixture of the colours corresponding to the positions of the slits. It is found that, if the positions of three slits be suitably chosen, any colour whatever can be matched by the mixture of light of these three wave-lengths taken in various proportions.

The three primary colours, by mixing which all the various colour sensations can be obtained, are violet near the Fraunhofer line G, green between E and F, and red between B and C.

If the intensities of the three primaries taken are violet 250, green 203, and red 100, the mixture produces the same sensation as white light straight from the arc.

398. The Young-Helmholtz Theory of Colour.—From the fact that any colour sensation could be produced by the mixture, in suitable proportions, of light of three given wave-lengths, Thomas Young was led to suppose that there existed three primary colour sensations, and Helmholtz has supposed that the reason for this is that the eye is furnished with three sets of nerves, one set which, when excited, gives the sensation of red, another of green, and the third of violet. When more than one set of nerves is excited, then a mixed sensation is produced, the character of which depends on the degree to which each set of nerves has been excited.

According to the Young-Helmholtz theory of vision it is supposed that each set of nerves, the red say, transmits the sensation of red to the brain, whatever the manner in which they may have been stimulated. Thus the red nerves are affected not only by red light but also, to a smaller extent, by light of other wave-lengths; the impression produced on the brain is, however, always that of red light.

It has been found possible, by studying the colour sensations of normal-eyed and of colour-blind persons, to draw three curves showing the sensitiveness of the three primary sets of nerves to stimulation by light of different wave-lengths. Such a curve is shown in Fig. 377, and was obtained by Koenig. It will be seen from this curve that the sensation of red can be stimulated by light of all wave-lengths, as is also very nearly the case with the green nerves. The violet nerves, however, are not at all affected by the red end of the spectrum.

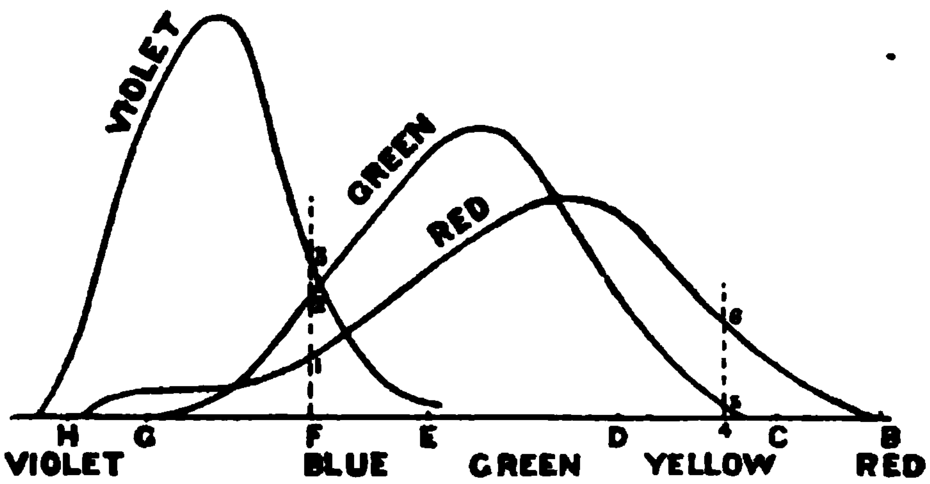


FIG. 377.

The above theory as to three sets of colour nerves accounts satisfactorily for the abnormal colour sensations of colour-blind persons. Thus a colour-blind person is one in which one (very rarely two) of the sets of colour nerves is missing. Thus a red colour-blind person is one in which the red nerves are insensitive. Hence such a person will only possess the violet and green sensations, and it will at once be perceived why the luminosity curve (Fig. 375) obtained by such an observer falls so much below the normal at the red end of the spectrum, for there the green and violet sensations are very weak.

In the case of a green colour-blind person, the green sensation is absent, and hence the curves in Fig. 377 show that for blue light, about half-way between E and F, the two remaining sensations are equally stimulated. Now equal stimulation of all three sensations corresponds, in the normal eye, to white, and in the green colour-blind, in the same way, equal stimulation of the violet and red sensations also corresponds to the sensation produced by white. Hence when blue light enters the eye of a green colour-blind person the impression produced is the same as that produced by white light. This fact is brought out very clearly if such a colour-blind person is shown a spectrum, for he will say that he sees red at one end and violet at the other, with a white band between.

399. Complementary Colours.—Two colours are said to be complementary when, if combined, they produce the sensation of white.

The complementary colours of the spectrum colours can be obtained by stopping the light of any given wave-length by means of an opaque rod and allowing the remaining colours to be combined by the lens L_3 (Fig. 374) on the screen at B. Then the colour seen at B will be the complementary of that which is removed from the spectrum by the interposition of the rod.

Since not only all the spectrum colours, but also the three primary colours taken in their proper proportions, produce the sensation of white, if one of the slits is closed the colour produced by the mixture of the remaining two will be complementary to the missing colour.

It is not necessary that either of the two complementaries should be a compound colour, for if a slit be placed in the blue near F, and another in the yellow between D and C, the mixture of the two simple colours transmitted will produce the sensation of white, and hence these two colours are complementary. The reason for the production of the sensation of white by the mixture of the above two colours can be seen by a study of the curves in Fig. 377. For it will be seen that the sum of the ordinates of the red and green curves, where cut by the dotted lines F, 1, 2, 3 and 4, 5, 6, are each equal to the ordinate, F3, of the violet curve. Hence the combined effect of these two lights is to excite all three primary sensations to an equal extent, that is, to produce the sensation of white.

Although a mixture of blue and yellow light produces the sensation of white, it is otherwise if we mix blue and yellow pigments, for in this case the result is a pigment which, when illuminated by white light, produces the sensation of green. The reason for this is that in the case of pigments the light which reaches the eye is white light which has been deprived of some of its components by absorption within the pigment. Thus a blue pigment will absorb all the colours except the blue and green, while a yellow pigment will absorb all but the red, yellow, and green. Now suppose we have a mixture of fine yellow and blue pigment particles illuminated by white light, the blue particles will absorb all the components of the white light except the blue and green, but will transmit these two colours. The yellow particles will absorb the blue but will also transmit the green. Thus all the components of the white light will be absorbed, by one or other of the two kinds of particles, except the green, and hence all the light which is transmitted or reflected from the pigment will be green. The truth of this explanation can be proved by painting a card with yellow pigment and holding it in a beam of light which has passed through a blue solution. Blue and green light will now fall on the yellow pigment, and of this the blue will be absorbed and the green will be reflected, so that the card appears green. In the same way a card painted blue, when illuminated by light obtained by passing white light through a yellow solution, will also

appear green, for of the incident yellow and green light the yellow will be absorbed by the pigment.

Experiments with pigments led to the conclusion that red, yellow, and blue were the three primary colours, for the red pigment will absorb the green which is transmitted by the other two, and so a neutral tint is produced. Thus when using pigments to examine the phenomena of colour great care must be taken, for in no case are pigment colours, even approximately, monochromatic ; and it must always be remembered that the colour of a pigment is obtained by the absorption of light of certain wave-lengths from the incident light.

CHAPTER X

POLARISATION AND DOUBLE REFRACTION

400. Light transmitted by Tourmaline—Polarisation.—We have supposed that light is due to a wave-motion in the ether, but have not yet considered whether the waves are longitudinal, such as is the case with sound-waves, or are transverse, *i.e.* whether the displacement, whatever its nature may be, which causes the sensation of light (and also, of course, of radiant heat) takes place normally to the wave-front or parallel to the wave-front. This question can be answered at once by means of an experiment made with two crystals of tourmaline.

If we take a slice of a crystal of tourmaline cut parallel to the crystallographic axis, and pass a ray of light through it, part of the light will pass through, and will, with most specimens of tourmaline, be coloured greenish owing to selective absorption within the crystal, otherwise to the eye the character of the light appears unaltered, and remains of the same intensity if the tourmaline plate is rotated. If the light which has passed through one tourmaline plate is allowed to fall on another, placed with its axis parallel to the first, the light will pass through the two; the only visible effect will be to *slightly* darken the greenish tint, the intensity being very slightly diminished by the second plate. If, however, the second plate is gradually rotated round an axis parallel to the light, so that the axes of the two crystals are inclined at a finite angle to one another, the intensity of the transmitted light will gradually diminish, till, when the axes are at right angles, none of the light which has passed through the first plate will pass through the second.

Hence the light which has passed through a plate of tourmaline has acquired properties which it did not before possess, in that it can no longer pass through a second plate of that substance when this plate is turned so that its axis is perpendicular to the axis of the first plate.

In order to see to what conclusions this experiment leads, let us consider an analogous problem. If we have a stretched string, we have seen that it is capable of two distinct modes of vibration, namely, a longitudinal vibration, in which the particles of the string move backwards and forwards in the direction of the length of the string, and a transverse vibration, in which the particles move in planes perpendicular to the length of the string. In the case of the string vibrating longi-

tudinally, the appearance of the string is the same on all sides, *i.e.* it remains stretched straight between its extremities. When it is vibrating transversely, however, its appearance is ordinarily different on different sides, since it vibrates in a single plane. Hence a string vibrating transversely has definite sides, so that, to define its motion with reference to the surrounding medium, we must state the plane, passing through the undisturbed position of the string, in which the vibration takes place. Another kind of transverse vibration of which a string is capable is that in which each particle describes a circle in a plane at right angles to the undisturbed position of the string. Suppose, then, that we cause a string to vibrate in this manner by attaching one end to a hook fixed at a little way from the centre of a rapidly rotating disc A (Fig. 378). The string will appear to swell out into something like the shape shown by the dotted lines. If now the string is passed through a narrow vertical slit D, since the motion of the string can then only take place up and down this slit, beyond D the motion will consist of transverse vibrations executed in a plane passing through the undisturbed position of the string and the slit, and by rotating the slit the plane in which the vibrations are taking place will also be rotated. Next, if a second slit is placed at F, and this slit is parallel to the first, the motion of the string, being parallel to this slit, will be unaffected. If, however, the first slit remaining vertical, the second slit F is turned out of the vertical, it will begin to interfere with the vibration of the cord, and when it is horizontal it will no longer allow any of the motion of the cord, which is in a verticle plane, to pass, and hence the portion of the cord between the second slit and B will remain at rest.

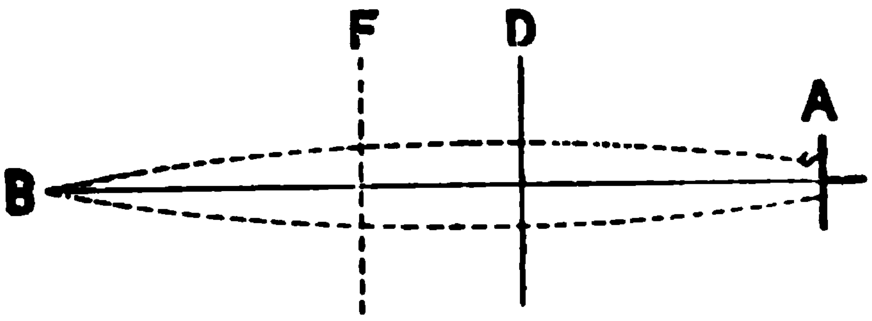


FIG. 378.

The experiment with the crossed tourmalines gives just such a result as the above, and so we conclude that the reason the light will not pass through the second tourmaline, when the axes are at right angles, is that during its passage through the first the light vibrations have acquired sides, or, in other words, they now occur in one plane, so that they are stopped by the second tourmaline, just as the transverse vibrations of the cord are stopped by the second slit after they have been confined to one plane by the first.

Since no such action could take place with the cord vibrating longitudinally, we conclude that the light vibrations are transverse.

Ordinary light, then, consists of transverse vibrations, and since when a *single* tourmaline is used the intensity of the transmitted light does not change as the tourmaline turns, the vibrations must take place

in all directions at right angles to the direction of the ray. After the passage of the light through the tourmaline, however, the transverse vibrations all take place parallel to some definite direction, and the ray is said to be *plane polarised*. Thus when a ray of light IO (Fig. 379) passes through the tourmaline plate AB, which is cut so that the axis of the crystal is parallel to AB, the transmitted light is plane polarised, *i.e.* the vibrations take place in one plane. As has been mentioned in § 366, we do not know for certain what is the nature of the wave-motion in the ether which we call light, and we cannot say whether the vibrations take place in the plane containing the axis of the crystal and the ray, or in the plane at right angles. Fresnel assumed in his theory that the vibrations take place in the plane containing the ray and parallel to the axis, as shown by the small double-headed arrows in the figure. M'Cullagh, on the other hand, considered that the vibrations take place in planes at right angles to the axis. For the purpose of explaining the phenomena we shall adopt Fresnel's hypothesis, merely remarking that according to the electro-magnetic theory of light there is something going on in both planes, in one an electric vibration, and in the other a magnetic vibration.

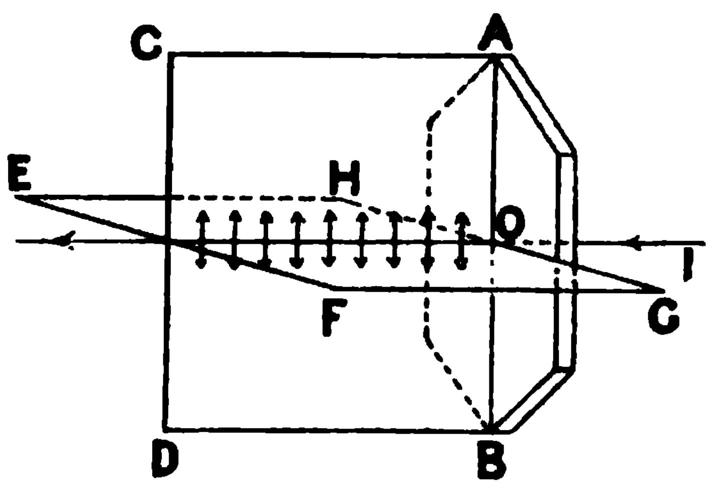


FIG. 379.

Hence we assume that the vibrations in the ether which produce light take place in the plane ABDC containing the axis. The plane FEHG, drawn through the direction of the ray and at right angles to the direction of vibration of the ether particles, is called the plane of *polarisation*.

Ordinary light is light in which, at any instant, the ether particles at a given point are vibrating in straight lines, but such that the direction of the vibrations changes from time to time. Since interference can be observed between rays which have traversed paths which differ by 400,000 wave-lengths, it follows that during the time taken for this number of vibrations, *i.e.* 10^{-9} second, the direction of the plane of polarisation does not appreciably change. A much slower change than this would, however, be imperceptible, since the eye is unable to appreciate periodic changes which take place in a time less than .05 of a second.

401. Double Refraction.—In treating of the refraction of light we have hitherto assumed that the media between which the light passes are both isotropic. We have now to consider the refraction of light when one of the media is æolotropic.

An æolotropic body which is transparent to light is the crystalline

calcium carbonate, called Iceland spar, this substance possessing different physical properties in different directions. It crystallises in various forms, but they all split most readily along certain planes which are always inclined to each other at fixed angles, so that by cleavage the crystals can always be reduced to the rhombohedral form shown in Fig. 380. The rhombohedron is bounded by six parallelograms, each of which has two acute angles of 78° 5′ and two obtuse angles of 101° 55′. Two of the solid angles, A and B (Fig. 380), are formed by three obtuse angles, the remaining four being formed by one obtuse and two acute angles. A line drawn through either A or B, so as to be equally inclined to the three sides or edges which meet at the corner, is called the axis of the crystal. We shall find that the axis has very distinct and special optical properties, and since these properties are unaltered if the length, breadth, or thickness of the crystal are altered, as far as such optical properties are concerned, the axis must be looked upon as simply a direction in the crystal, so that all lines parallel to the crystallographic axis are optical axes.

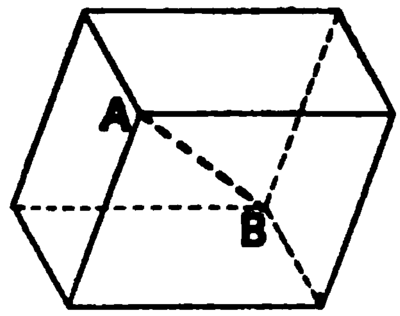


FIG. 380.

If a ray of light is incident normally on one of the faces of a rhombohedron of Iceland spar, part of the light will pass straight through along POO′ (Fig. 381), just as would happen in the case of an isotropic body. Part of the light will, however, be refracted and travel along PEE′. Hence in this case there are two refracted rays corresponding to a single incident ray. This phenomenon is spoken of as *double refraction*.

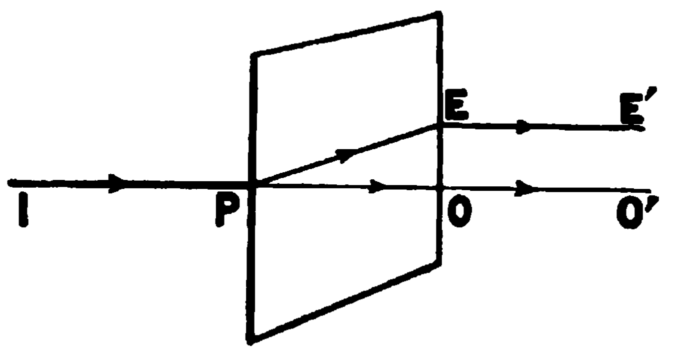


FIG. 381.

If the Iceland spar is turned round the line PO as an axis, and the transmitted light is received on a screen, the spot of light corresponding to the refracted ray POO′ will remain stationary, while that due to the ray PEE′ will rotate round the other as a centre, the line joining the two images being always parallel to the shorter diagonal of the parallelogram[1] which constitutes the face of the crystal on which the light is incident normally.

If the angle of incidence is not zero, then in general there will be two refracted rays, but while one of them obeys Snell's law in that the ratio of the sine of the angle of incidence to the sine of the angle of refraction is constant, whatever the direction of the incident ray, for the other ray this ratio is different for rays incident in different directions.

The refracted ray PO, which obeys the ordinary law of refraction, is

[1] This is only true if the edges of the face are all of equal length. If this condition is not fulfilled, such an equilateral parallelogram must be marked out on the face.

called the *ordinary ray*, while the other ray PE, which does not obey this law, is called the *extraordinary ray*.

If the light which has been transmitted through the crystal of Iceland spar is allowed to fall on a plate of tourmaline, and if this plate is rotated, it will be found that in some positions only the ordinary ray is transmitted, and in others only the extraordinary ray. Since the positions for which the ordinary ray is extinguished are at right angles to the positions in which the extraordinary ray is extinguished, this experiment shows that both the ordinary and the extraordinary rays are plane polarised, and that the planes of polarisation are at right angles to one another.

If in the case of a doubly refracting crystal a plane be drawn perpendicular to the face on which the light is incident, and so as to contain the optic axis, this plane is called the *principal plane* for the given face.

When the axis of the tourmaline plate is parallel to the principal plane of the spar, the ordinary ray is cut off by the tourmaline, which shows that the ordinary ray is polarised in a plane parallel to the axis of the tourmaline, *i.e.* parallel to the principal plane of the spar. When the axis of the tourmaline is at right angles to the principal plane the extraordinary ray is cut off, so that this ray is polarised in a plane perpendicular to the principal plane.

Thus suppose ABCD (Fig. 382) represents the cross section, taken perpendicular to the faces AD and BC, of a plate of Iceland spar cut through the corners of the rhombohedron, so that the line AX, or any line parallel to this line, is the optic axis, the plane of the paper will be the principal plane of the plate for light incident on the faces AD or BC, since it is perpendicular to these faces and also contains the axis. Then the plane of the paper is the plane of polarisation for the ordinary ray, so that in this ray the ether vibrations take place along straight lines perpendicular to the plane of the paper, while the extraordinary ray is polarised in a plane at right angles to the paper, and the vibrations take place in the plane of the paper. The directions of the two rays for various angles of incidence are shown in the figure.

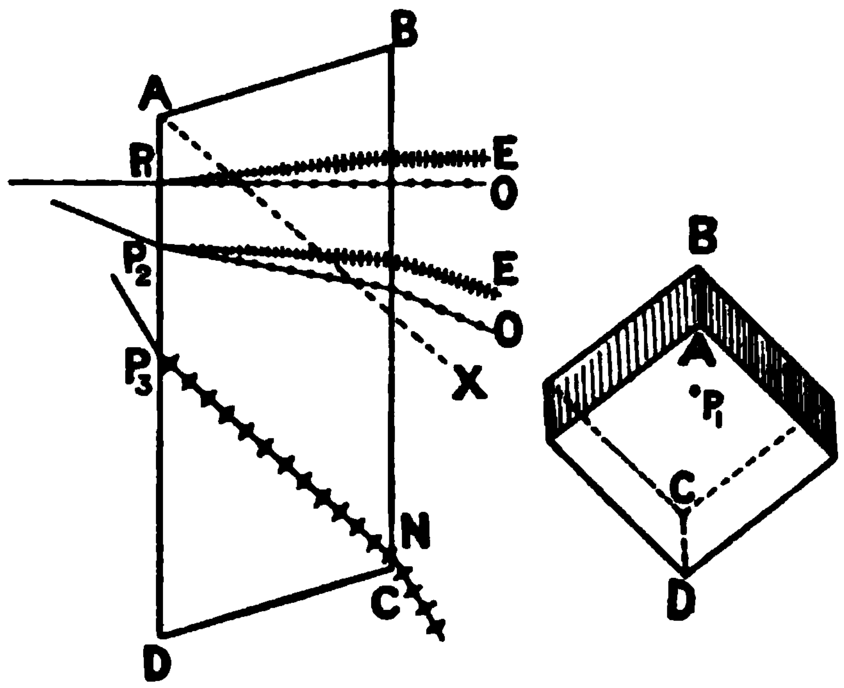


FIG. 382.

When the refracted ray is parallel to the axis, as at P_3N, there is only

a single refracted ray, and this ray obeys Snell's law, *i.e.* the extraordinary ray for this angle of incidence coincides with the ordinary ray.

If ABC (Fig. 383) represents a section of a prism of Iceland spar cut so that the refracting edge is parallel to the optic axis, the plane of the paper is perpendicular to the principal plane, and hence the ordinary ray POO′ is polarised in a plane perpendicular to the paper, and the extraordinary ray is polarised in the plane of the paper. By setting the prism so that first the ordinary and then the extraordinary ray is at minimum deviation, the refractive index for the ordinary ray (a constant) can be determined, as also the refractive index for the extraordinary ray in a plane at right angles to the axis. The refractive index for the *D* line for the ordinary ray is 1.658, while that for the extraordinary ray in a plane perpendicular to the axis, *i.e.* in a plane at right angles to the principal plane, is 1.486.

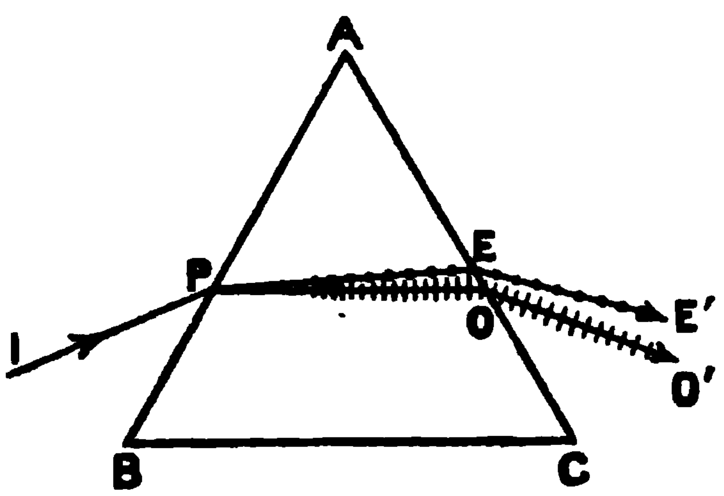


FIG. 383.

Let ABCD (Fig. 384) represent the face of a section of a crystal of Iceland spar cut parallel to the optic axis XX′, then if a ray of ordinary light is incident normally at O, it will be split up by its passage through the crystal into two rays of equal intensity; one of these, the ordinary, is polarised in a plane perpendicular to the paper passing through XX′ (the principal plane, since it is perpendicular to the face ABCD, and contains the axis XX′), and the other, the extraordinary, polarised in a plane perpendicular to the paper and passing through YY′.

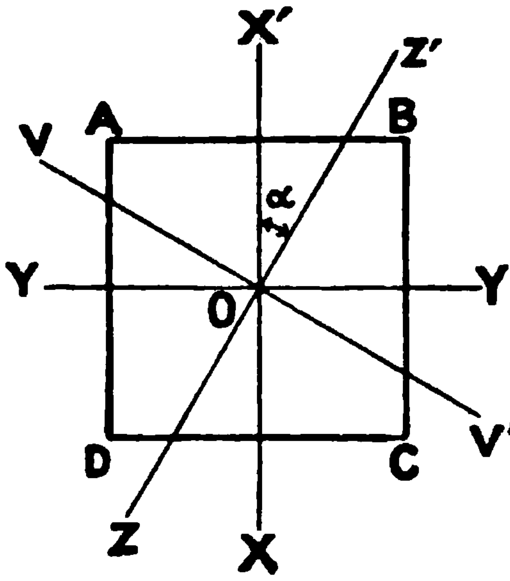


FIG. 384.

Next suppose that, in place of the incident light being unpolarised light, we use plane polarised light, say by allowing the light to pass through a tourmaline plate. If the axis of the tourmaline is parallel to XX′, the incident light is polarised in the plane passing through YY′ and the vibrations of the ether take place parallel to XX′. Now the vibrations of the ordinary ray take place parallel to YY′, since the ordinary ray is polarised in the plane XX′, so that the motion of the ether particles in the incident light is at right angles to the direction of motion of the ether particles in the ordinary ray. Hence the motion of the incident waves has no component in the direction in which the vibrations of the ordinary ray take place, and so cannot give rise to such vibrations. There will therefore be no ordinary

ray when the incident ray is polarised in the plane XX′. In the same way, there will be no extraordinary ray when the incident ray is polarised parallel to the plane YY′.

Next suppose that the incident light is polarised in some intermediate direction, say parallel to the plane ZZ′, which makes an angle a with the optic axis XX′, the ether vibrations taking place along VV′, which is inclined at an angle $90° - a$ to the optic axis. The incident vibration will now have a component along both XX′ and YY′, so that there will be both an ordinary and an extraordinary ray. In order to find the relative intensities of these rays, we may resolve the incident vibration along XX′ and YY′. Since the motion of the ether particles is along VV′, if A is the amplitude of the incident vibration, the amplitude of the vibration parallel to XX′ will be A cos VOX′ = A cos $(90° - a)$ = A sin a, while the amplitude of the vibration parallel to YY′ will be A cos VOY = A cos a.

Thus, since the energy of a vibratory motion is proportional to the square of the amplitude, we have that the intensity of the incident light is A^2, that of the ordinary ray $A^2 \cos^2 a$, and that of the extraordinary ray is $A^2 \sin^2 a$. Since

$$\begin{aligned} A^2 \cos^2 a + A^2 \sin^2 a &= A^2(\cos^2 a + \sin^2 a) \\ &= A^2, \end{aligned}$$

we see that the sum of the intensities of the ordinary and the extraordinary rays is equal to the intensity of the incident light.

If a ray of light is allowed to pass through two plates of Iceland spar it will easily be seen, from what has already been said, that if the principal planes of the two plates are parallel there will be one ordinary ray and one extraordinary, the ordinary ray in one crystal becoming the ordinary ray in the other, &c. If the principal planes are at right angles, the ordinary ray in one crystal will become the extraordinary ray in the other, there being, as before, only two rays. When, however, the principal planes are inclined at an angle between 0° and 90°, there will be four transmitted rays, since the ordinary and extraordinary rays in the first plate will each give rise to an ordinary and extraordinary ray in the second plate, the intensities of the rays varying with the angle between the principal planes.

Double refraction is exhibited by all crystalline bodies except those which crystallise in the cubic system. In every case the ordinary and extraordinary rays are plane polarised in planes at right angles to one another. In the case of tourmaline, the incident light is split up into an ordinary and an extraordinary ray, but this substance exerts a powerful selective absorption on light polarised in a plane containing the axis, *i.e.* the ordinary ray, so that with moderate thicknesses of the crystal only the extraordinary ray can pass.

402. Interference of Polarised Light.—By placing two plates of tourmaline, so that each is traversed by the rays passing through one half

of Fresnel's biprism, the interference of polarised light can be studied in the manner described in § 373. In this way it is found that two rays of light polarised in planes at right angles do not produce interference under circumstances in which two rays of ordinary light would interfere. Two rays polarised in the same plane do, however, interfere like two rays of ordinary light. The fact that rays polarised in planes at right angles do not interfere is a further proof that the direction of vibration in the case of light is transverse to the direction of propagation.

403*. Uniaxal and Biaxal Crystals.—We have, when speaking of double refraction in Iceland spar, mentioned that when a ray of light traverses the spar parallel to the optic axis there is only a single refracted ray. In Iceland spar there is only one direction in which this takes place, and therefore there is only one optic axis. Doubly refracting crystals which have only one optic axis are called *uniaxal* crystals. In other crystals there are two axes along which there is only a single refracted ray, and these are called *biaxal crystals.*

404*. Wave-Surface in Uniaxal Crystals.—If we have a disturbance produced at a point within an isotropic medium, the wave-surface at any moment will be a sphere with the point of disturbance as the centre, for the velocity of light being the same in all directions, the disturbance which originates at any point will in a given time spread to an equal distance in all directions. If, however, the body is not isotropic, and the velocity of light is different in different directions, the disturbance will, in a given time, travel further in some directions than in others, and so the wave-surface will no longer be a sphere.

Now, in the case of a crystal, the velocity with which light travels is not the same in all directions; and since there are in general two refracted rays there must be two wave-fronts. For the ordinary ray the refractive index is constant, and therefore the velocity of the ordinary ray in the crystal is constant, for we have seen in § 366 that the refractive index is equal to the ratio of the velocities in the two media (air-crystal). As the refractive index for the extraordinary ray varies with the direction of the ray within the crystal, the velocity with which the extraordinary ray travels must depend on the direction of the ray in the crystal. The velocity in the case of the ordinary ray being constant, just as in isotropic bodies, the ordinary wave-surface must be a sphere. Huyghens, who first considered the subject, assumed that in uniaxal crystals the extraordinary wave-surface was a *spheroid* or *ellipsoid of revolution*, that is, the figure obtained by rotating an ellipse about one of its diameters, AB or CD (Fig. 385), and then verified the accuracy of this assumption experimentally. The axis about which the ellipse is rotated to form the extraordinary wave-surface coincides with the optic axis of the crystal. Hence the complete wave-surface for a disturbance

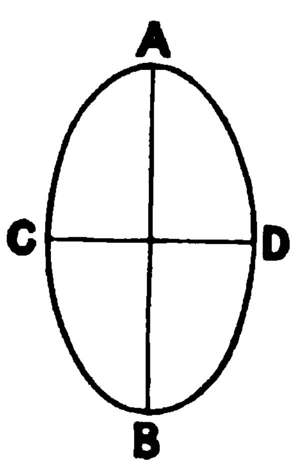
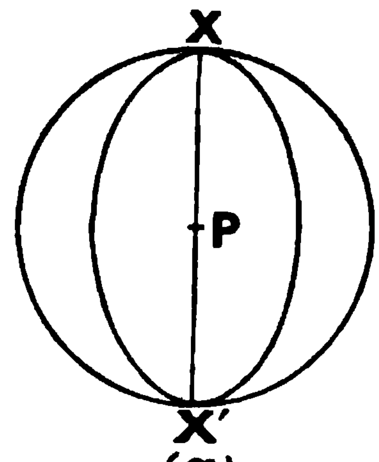


FIG. 385.

originating at a point within a uniaxal crystal consists of a sphere and a spheroid, both having their centres at the point; the axis of the spheroid being parallel to the optic axis of the crystal. Further, since when the ray travels in the crystal in a direction parallel to the optic axis, there is only one refracted ray, *i.e.* the ordinary and extraordinary rays travel with the same velocity, the sphere and spheroid must touch one another at the extremities of the optic axis.

Two cases, however, may arise. In the first place, the extraordinary ray may be more refracted than the ordinary ray, so that the velocity of the extraordinary ray is *less* than that of the ordinary ray. In this case, Fig. 386 (*a*), the spheroid lies within the sphere, touching it on the optic axis XX′. In the other case the velocity of the extraordinary ray may be *greater* than that of the ordinary ray, so that the spheroid lies outside the sphere, Fig. 386 (*b*), again touching the circle on the optic axis XX′.



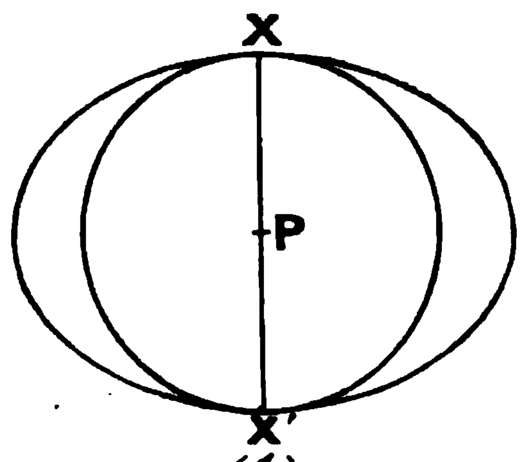


FIG. 386.

Uniaxal crystals in which the wave-surface is like Fig. 386 (*a*), and in which, except along the optic axis, the ordinary ray travels faster than the extraordinary ray, are called *positive* crystals. Quartz and ice are positive crystals.

Uniaxal crystals, in which the wave-surface is like Fig. 386 (*b*), are called *negative* crystals, and to this class belong Iceland spar and tourmaline.

405*. Huyghens's Construction for the Directions of the Refracted Rays in a Uniaxal Crystal.—Suppose we require to find the directions of the refracted rays in the case of Iceland spar. The spar being a negative crystal, the wave-surface is like Fig. 386 (*b*).

Let IQ or I′P (Fig. 387) be the direction of the light incident on the face of the crystal, and let the optic axis XQ lie in the plane of the paper, so that the paper is the principal plane for the face. Then QM is the wave-front of the incident wave.

When the wave reaches Q, we may consider that this point becomes a centre of disturbance within the crystal. If it takes the wave a time t to travel from M to P, and we describe the wave-surfaces in the crystal about the point Q for a time t after the disturbance reaches Q, these wave-surfaces will represent the positions of the wave in the crystal when the wave in the air reaches P. Hence, if from P we draw $\overline{PO}$ and $\overline{PE}$ tangents to the ordinary and extraordinary wave-surfaces respectively,

PO will represent the ordinary wave-front in the crystal and PE the extraordinary wave-front, and the line QO will represent the direction of the ordinary ray and QE the direction of the extraordinary ray.

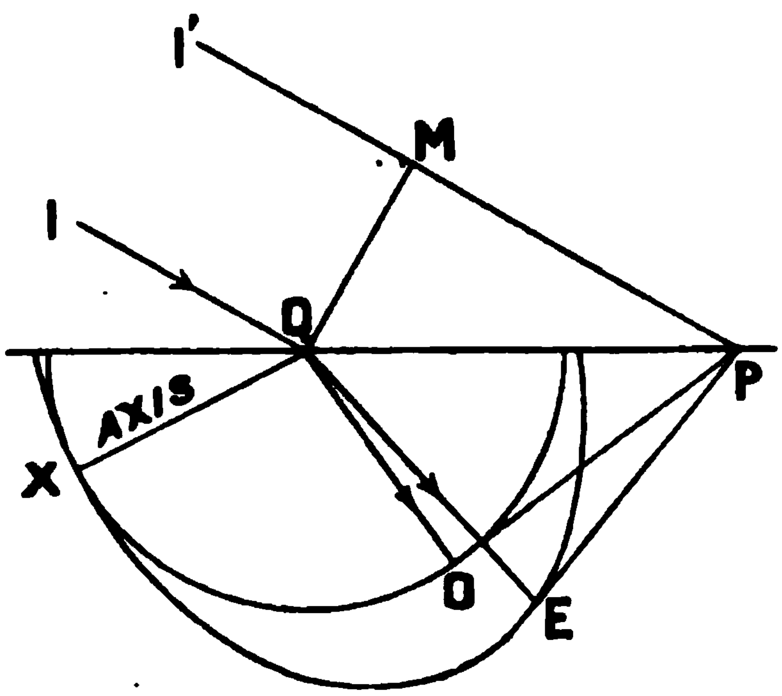


FIG. 387.

If the plane of the paper had not been a principal plane, we should have had to draw through P a plane perpendicular to the plane of incidence to touch the spheroid, and it would not have touched it at a point in the plane of the paper; so that the extraordinary ray would not be in the plane of incidence, and thus would not have obeyed the first law of refraction as given in § 341.

Two particular cases are worth examining: first, when the optic axis is parallel to the face of the crystal and perpendicular to the plane of incidence; and second, when the optic axis is parallel to the face and also parallel to the plane of incidence.

In the first case (Fig. 388) the optic axis is perpendicular to the plane of the paper, and hence the sections of the wave-surfaces consist of two circles, the inner one, since the crystal is negative, corresponding to the ordinary ray. The reason the section of the extraordinary wave-surface is a circle is that this surface is obtained by rotating an ellipse about the optic axis, so that all sections perpendicular to the axis must be circles.

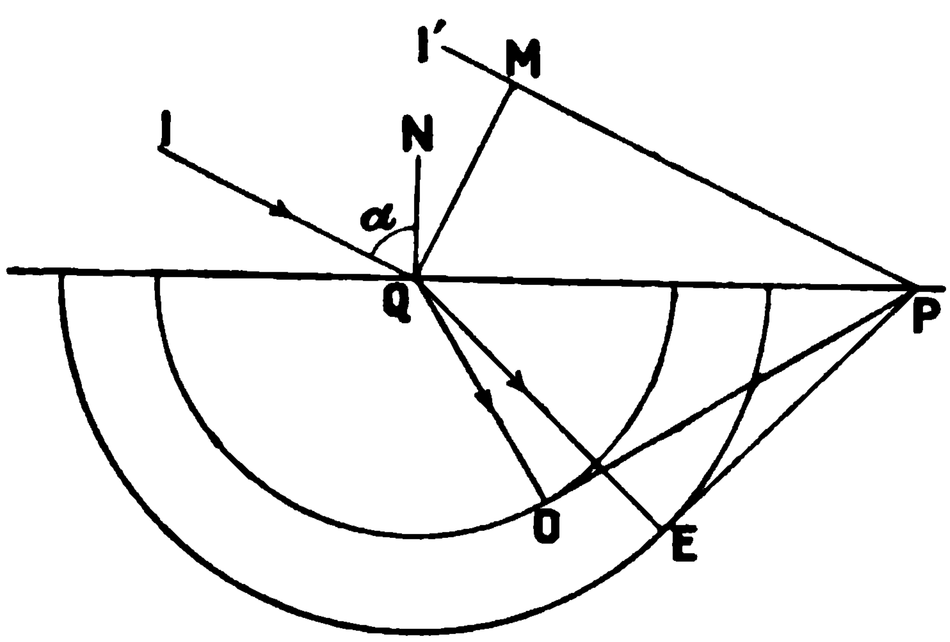


FIG. 388.

If a is the velocity of the ordinary ray and b the velocity of the extraordinary ray in a plane at right angles to the optic axis, the radius of the spherical wave-surface in the crystal being taken as a, the major axis of the spheroid will be b. Hence in Fig. 388, if $\overline{QO}$ is a, $\overline{QE}$ will be equal to b. If the velocity of light in air is c, the refractive index for the ordinary ray is c/a, and that for the ordinary ray in a plane at right angles to the optic axis is c/b. Now b or $\overline{QE}$ is

constant for all angles of incidence in a plane at right angles to the optic axis, and hence the extraordinary refractive index is constant in this plane, and the extraordinary ray obeys the ordinary laws of refraction. By cutting a prism of Iceland spar with its refracting edge parallel to the optic axis two refracted rays will be obtained, and the refractive index (c/a and c/b) corresponding to each of these can be measured. In this way it can be proved that the extraordinary refractive index (c/b) in a plane at right angles to the optic axis is constant. Hence b or $\overline{QE}$ must be constant, and so it is proved that the section of the extraordinary wave-surface perpendicular to the axis is a circle.

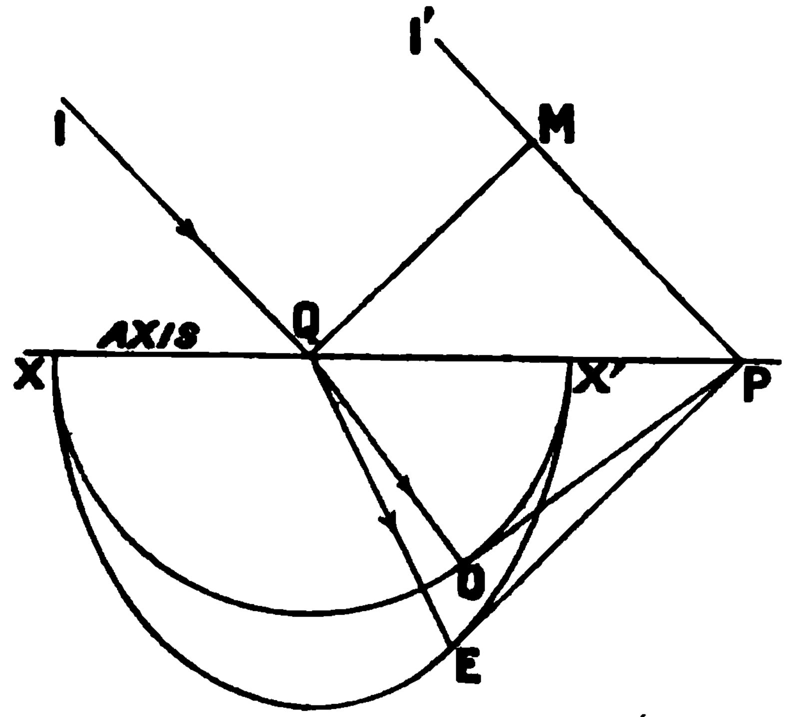


FIG. 389.

The construction for finding the directions of the refracted rays when the optic axis is parallel to the face of the crystal, and in the plane of incidence, is shown in Fig. 389.

406. Nicol's Prism.—As a means of obtaining plane polarised light, a tourmaline plate is, for many purposes, unsuited, for, as has been mentioned, the light transmitted by tourmaline is coloured green. Since, when a beam of light is passed through a crystal of Iceland spar, two refracted beams are obtained, each of which is plane polarised, but in planes at right angles, if by any means we could intercept one of these refracted beams, the other would give us plane polarised light. Since the angular separation between the ordinary and extraordinary rays is not very great, it is not possible to stop one of the beams with a screen, unless only a very narrow beam is employed, or we use a very thick crystal.

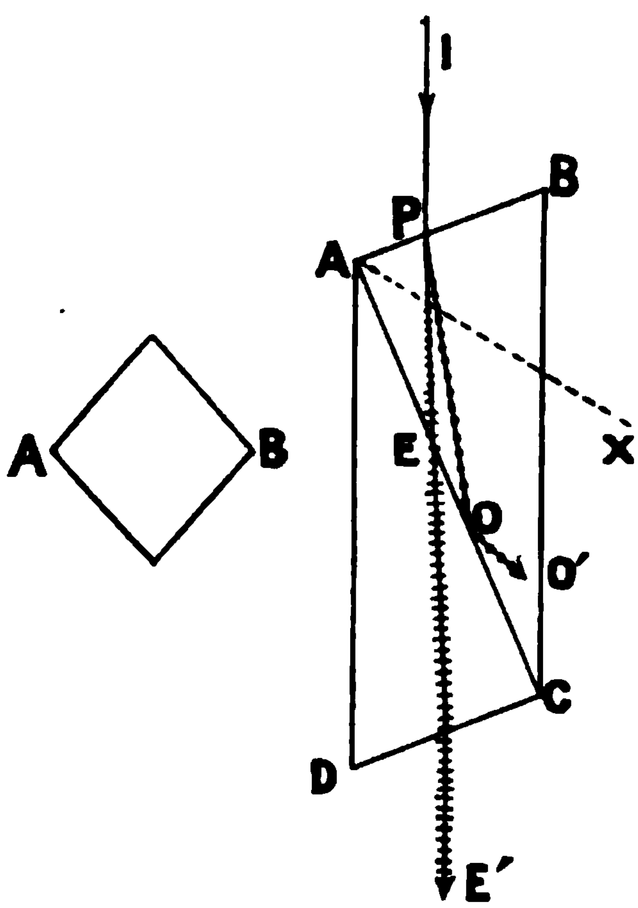


FIG. 390.

The most convenient method of getting rid of one of the rays is to make use of total internal reflection for this purpose. A rhomb of Iceland spar is taken and cut in two by a plane, AC (Fig. 390), perpendicular to the principal plane for the face AB. The two

surfaces are then polished and cemented together in their original position by means of a thin film of Canada balsam.

Now the refractive index of Canada balsam (1.55) is greater than the minimum value for the extraordinary ray (1.486) in Iceland spar, and less than that for the ordinary ray (1.658). As total reflection can only occur when light is passing from a media of greater to one of less refractive index, we can never get total reflection in the case of the extraordinary ray when passing from spar to balsam, so long as the ray passes in such a direction that the refractive index is less than 1.55. In the case of the ordinary ray, however, if the incidence is sufficiently oblique we shall obtain total reflection. Hence if the plane AC is suitably inclined, the ordinary ray, PO, will be incident on the surface AC at an angle greater than the critical angle, and will therefore be totally reflected along OO′, while the extraordinary ray, PEE′, will pass through the prism.

The light transmitted by such a rhomb of Iceland spar, which is called a Nicol's prism, will therefore be plane polarised, and since it is the extraordinary ray which is transmitted, the plane of polarisation is perpendicular to the principal plane, *i.e.* is a plane perpendicular to the paper in Fig. 390.

A Nicol's prism may be used, not only for producing plane polarised light, when it is called a polariser, but also for detecting whether light is plane polarised, and, if so, determine the plane in which it is polarised, when it is said to be used as an analyser.

If the light incident on the Nicol is unpolarised, then the intensity of the transmitted light will remain the same when the Nicol is rotated round the light ray as an axis, the intensity of the transmitted light being practically half that of the incident light. There is, however, a very slight loss due to reflection at E (Fig. 390), and where the ray leaves the crystal.

If the incident light is plane polarised, the intensity of the transmitted light varies as the analyser is rotated. When the principal plane of the Nicol is parallel to the plane of polarisation of the incident ray, then (§ 401) there will be only an ordinary ray in the spar, and this ray is totally reflected, so that no light will be transmitted. When the principal plane of the Nicol is perpendicular to the plane of polarisation of the incident light, only an extraordinary ray will be produced in the spar, and this will be transmitted undiminished, so that in this case the intensity of the transmitted light is equal to that of the incident light. If the principal plane of the Nicol is inclined at an angle α to the plane of polarisation, it can be shown, exactly as in § 401, that the intensity of the extraordinary ray, and hence that of the transmitted light, is $I \sin^2 \alpha$, where I is the intensity of the incident light. Thus when $\alpha = 0$ or $180°$ the intensity of the transmitted light is zero, and when $\alpha = 90°$ or $270°$ the intensity of the transmitted light is I.

407. Polarisation by Reflection.—If the light reflected from a non-metallic surface, such as glass, is examined with an analysing Nicol,

it will be found that as the Nicol is rotated the intensity of the transmitted light varies. For a certain angle of incidence there is no light transmitted by the Nicol when its principal plane is parallel to the plane of incidence of the reflected light, while when the principal plane of the Nicol is perpendicular to the plane of incidence, the transmitted light is equal in intensity to the reflected light before it passes through the Nicol. This shows that, for this angle of incidence, the reflected ray is completely plane polarised in the plane of incidence. For all other angles of incidence the reflected ray is only partly polarised, *i.e.* consists of a mixture of ordinary unpolarised light with light which is polarised in the plane of incidence. The angle of incidence, for which the reflected beam is completely plane polarised, is called the polarising angle for the reflecting substance.

If, instead of consisting of ordinary light, the incident ray is plane polarised, and is incident at the polarising angle, then when the incident ray is polarised in the plane of incidence, *i.e.* the vibrations of the ether are taking place perpendicular to the plane of incidence, the light will be reflected. If, however, the incident ray is polarised in a plane perpendicular to the plane of incidence, so that the ether vibrations take place in this plane, none of the light will be reflected, but it will all be refracted into the reflecting substance.

If the incident light is polarised in intermediate planes, the reflected light will gradually increase in intensity as the plane of polarisation changes from the position in which it is perpendicular to the plane of incidence, to that in which it is parallel to the plane of incidence.

Owing to polarisation by reflection, a glass plate can be used both as a polariser and as an analyser.

Suppose a ray of light IO (Fig. 391) is incident on a plate of glass A, at the polarising angle ϕ, which for ordinary glass is about 56°. The reflected ray, OP, will be polarised in the plane of incidence, that is, in the plane of the paper. If this reflected ray is received on a second glass mirror, B, also at the polarising angle, then if the plane of incidence on B is parallel to the plane of incidence on A, as is shown at (*a*) and (*b*), the polarised light will be reflected along PR. If now the mirror B is rotated round an axis parallel to OP, the angle of incidence will remain the same, viz. equal to the polarising angle, but the intensity of the reflected

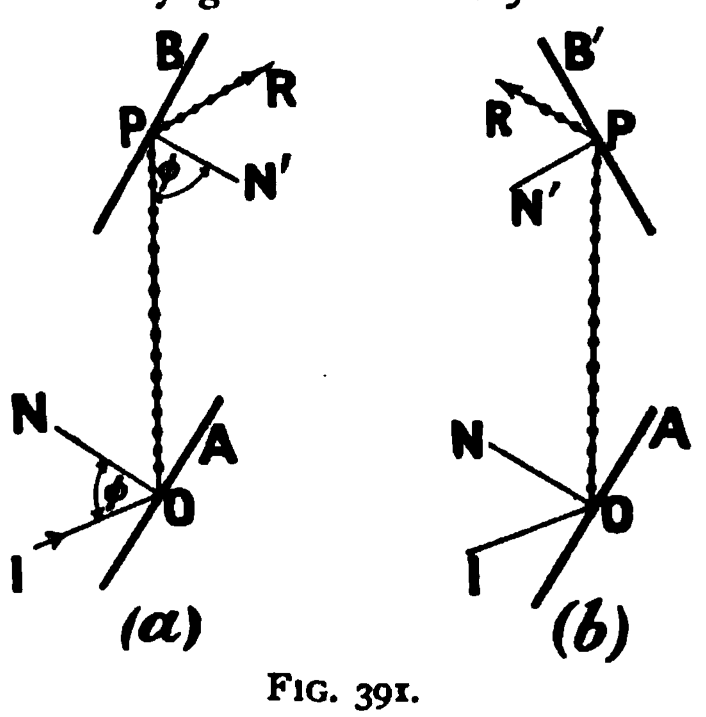


FIG. 391.

ray will diminish until, when the plane of incidence, which is of course the plane passing through P and O, and containing the normal PN′, is

perpendicular to the plane of the paper, there will be no reflected ray. Of course the reflected ray will again be of maximum intensity when the mirror has been turned through 180° into the position B′, and zero when it has been turned through 270°. Thus the one glass plate, A, has acted as a polariser and the other, B, as an analyser. This is the principle of Biot's and of Norrenberg's polariscopes.

408. Brewster's Law.—Sir David Brewster, having made an extensive series of experiments on the angle of polarisation for different substances, found that the tangent of the angle of polarisation is equal to the refractive index of the substance,[1] or if ϕ is the angle of polarisation and μ the refractive index

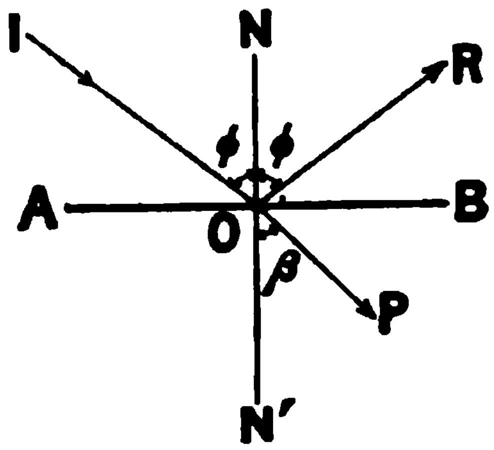


FIG. 392.

$$\mu = \tan\phi.$$

The geometrical interpretation of Brewster's law is very interesting. Let IO (Fig. 392) be a ray of light incident on a reflecting surface at the polarising angle ϕ, and OR and OP be the direction of the reflected and refracted rays respectively. If β is the angle of refraction N′OP, we have the following relations :—

By Snell's law $$\frac{\sin\phi}{\sin\beta} = \mu.$$

By the law of reflection the angle NOR $= \phi$.

But by Brewster's law $$\tan\phi = \frac{\sin\phi}{\cos\phi} = \mu.$$

$$\therefore \quad \frac{\sin\phi}{\cos\phi} = \frac{\sin\phi}{\sin\beta},$$

or $$\sin\beta = \cos\phi.$$

But $$\sin\beta = \cos(90° - \beta).$$

$$\therefore \quad \cos(90° - \beta) = \cos\phi.$$

or $$90° - \beta = \phi.$$

Hence the angle POB, which is equal to $90° - \beta$, is equal to ϕ, or to the angle NOR. Adding the angle ROB to each, we get that the angle POR is equal to the angle BON. Hence, since the angle BON is a right angle, the angle POR must also be a right angle ; that is, when the angle

[1] More recent observations by Jamin have shown that Brewster's law is only exact for substances for which μ is about 1.46. For substances of refracted index differing much from this value, the reflected beam is never entirely *plane* polarised, but for an angle of incidence given by the relation $\tan\phi = \mu$ the quantity of plane polarised light is a maximum.

of incidence is the polarising angle, the reflected ray OR is at right angles to the refracted ray OP.

Since the angle of polarisation depends on the refractive index, and this differs for light of different colours, the angle of polarisation will be different for different colours.

When the incident ray is plane polarised in the plane of incidence, the vibrations of the ether particles take place in straight lines parallel to

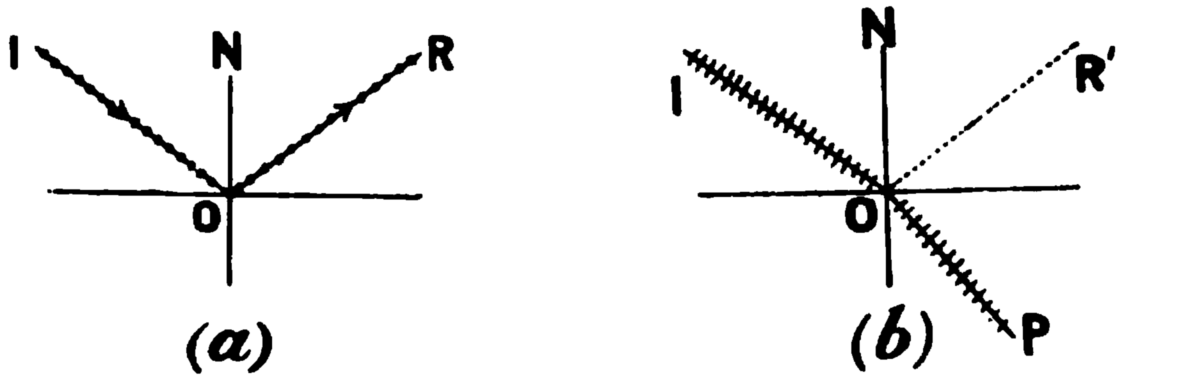


FIG. 393.

the reflecting surface as shown at (*a*), Fig. 393, and the vibrations in the reflected beam are also parallel to the reflecting surface, and to those in the incident ray.

When the incident ray is polarised perpendicular to the plane of incidence, the vibrations of the ether particles take place in the directions shown by the cross lines at (*b*), Fig. 393. Since the reflected ray OR′ is very nearly at right angles to the incident ray, the lines along which the ether particles vibrate in the incident ray are very nearly parallel to the direction of the reflected ray, and we should therefore expect that they would not produce any transverse disturbance along OR′; in other words, that there would be no reflected ray.

409*. Polarisation produced by Crystalline Media.—Suppose O, Fig. 394 (*a*), to represent the undisturbed position of a particle in an

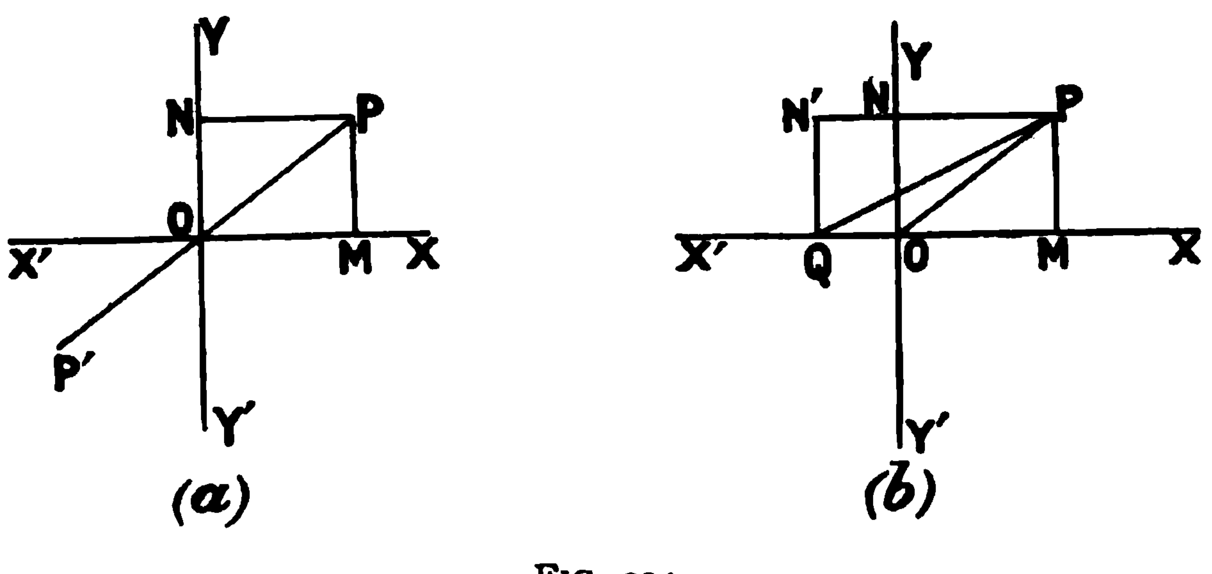


FIG. 394.

isotropic medium, and that the particle is displaced to a position P. Since the medium is by supposition isotropic, the elasticity must be the same in all directions.

Hence if OX, OY are two axes of reference, the elasticity, e, of the medium must be the same along OX as along OY. Since the elasticity is equal to *stress/strain*, the stress or restoring force called into play by a strain or displacement d in any given direction must be ed. Hence the restoring force in the direction PN is $e.\overline{PN}$, and that in the direction PM is $e.\overline{PM}$. The resultant of these two forces will therefore, by the parallelogram of forces (§ 66), lie along PO; that is, in the case of an isotropic body the restoring force acts along the direction of the displacement, and the particle P when released will vibrate along the line POP′.

In the case, however, of a crystal, where the elasticity is different in different directions, this will not be the case. For suppose that the elasticity (E) in the direction OX, Fig. 394 (b), is greater than that (e) in the direction OY. Then when the particle is displaced to P the restoring force along PN will be $E.\overline{PN}$, and that along PM will be $e.\overline{PM}$. To find the direction of the resultant force acting on P we draw a parallelogram (rectangle in this case, since the two components are at right angles), of which the sides are proportional to the two restoring forces. Hence if $\overline{MQ}$ is taken such that $\overline{MQ} = \frac{E}{e}\overline{PN}$, then $\frac{\overline{MQ}}{\overline{PM}} = \frac{E.\overline{PN}}{e.\overline{PM}}$, and the diagonal $\overline{PQ}$ will represent the direction of the restoring force acting on P. Since this force does not act through O, the particle, when set free, will not vibrate backwards and forwards through O, as it did in the case of the isotropic body, but will have its direction of motion gradually changed, till it finally takes place backwards and forwards along XX′.

If the displacement takes place either along OX or OY, the restoring force acts through O and the particle will continue to vibrate along XX′ or YY′, as the case may be. We can thus understand how it is that, if in a crystal the ether has different elasticities in different directions, the vibrations always take place in two planes that are at right angles to one another.

410. Double Refraction produced in Isotropic Bodies by Strain.—It is possible to render an isotropic body temporarily doubly refracting by subjecting it to a stress. This phenomenon can be examined by means of a bar of glass AB (Fig. 395) which is held in a metal frame, so that by screwing down the screw C the bar can be bent. If the bar is placed between crossed Nicols, so that the length of the bar is inclined at 45° to the principal planes of the Nicols, it will produce no effect so long as it is unstrained. On bending the bar, however, the light which passes through the parts of the bar above and below the median line will be able to pass through the analysing Nicol, while the central line, shown dotted in the figure, remains dark as before.

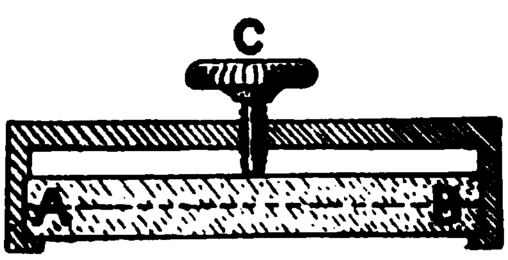


FIG. 395.

When the bar is bent, the part above the dotted line is compressed and the part below is extended, while the central part is unstrained. Since the central part is unstrained, it produces no effect on the plane of polarisation of the light which passes through it, and this light is entirely cut off by the analyser. The strained parts, on the other hand, have become doubly refracting, and the incident plane polarised light is partly decomposed into light polarised parallel and at right angles to the length of the bar, that is, at 45° to the principal plane of the analyser, and so will be able to pass through.

This method of placing a body between crossed Nicols, and seeing whether any light is then able to traverse the analyser, is a very delicate method of testing whether a transparent body is in a state of strain, and we shall see that under certain conditions even liquids may become doubly refracting due to strain.

411. Rotation of the Plane of Polarisation.—If two Nicol's prisms P and A (Fig. 396) are placed in line, the one, P, to act as a polariser, and the other, A, to act as an analyser, and if A is rotated till its principal plane is at right angles to that of P, on allowing a beam of parallel monochromatic light to fall on P, none of this light will pass A, for the light will be plane polarised in a plane at right angles to the principal plane of P, and hence when it falls on A there will only be an ordinary ray, since the incident light is polarised in the principal plane of this Nicol; and as this ray is stopped by total reflection, no light will be transmitted through A. If now a plate of quartz, Q, cut with the faces perpendicular to the optic axis, is placed between the Nicols, so that the light traverses the plate parallel to the axis, it will in general be found that some light is now transmitted by the analysing Nicol A. By rotating the analysing Nicol a position can, however, be found such that no light is again transmitted, and rotating the quartz plate round its axis produces no change. This experiment shows, in the first place, that when plane polarised light is passed through a plate of quartz parallel to the axis it remains plane polarised; and, in the second place, that the plane in which the light is polarised is *rotated* by the passage through the quartz. A body, such as quartz, which has this property of rotating the plane of polarisation of plane polarised light, is said to be *optically active.*

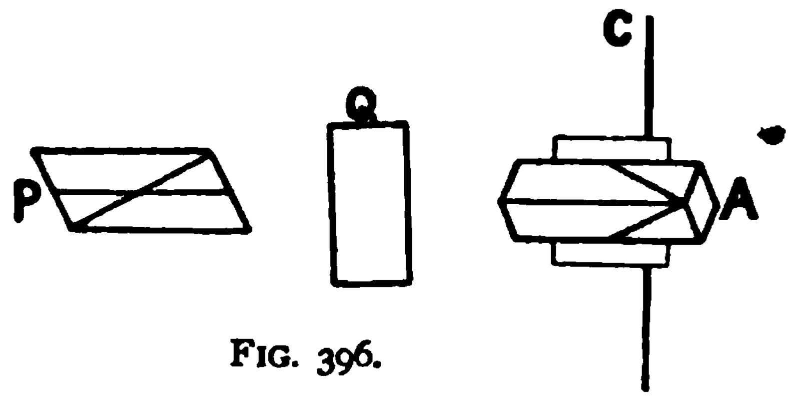


FIG. 396.

The amount of the rotation is proportional to the thickness of the substance traversed by the light, and varies with the wave-length of the light and the nature and temperature of the substance. While some bodies rotate the plane of polarisation in one direction, others rotate it in the reverse direction. If, looking in the direction in which the light is

travelling, the plane of polarisation is rotated in the same direction as that in which the hands of a clock appear to move when we look at the face, the rotation is said to be right-handed or positive; if the rotation takes place in the opposite direction, it is said to be left-handed or negative.

The rotation in some samples of quartz is to the right, and in others to the left, the amount of the rotation produced by equal thicknesses being the same in both cases. Some liquids, such as turpentine, are also optically active, as well as solutions of some substances such as sugar and quinine.

The rotation varies very nearly inversely as the square of the wave-length, so that, if white light is used, when the analysing Nicol is set to extinguish light of any given wave-length, the light of the other wave-lengths will be transmitted. If the transmitted light is examined by means of a spectroscope, the spectrum will be seen crossed by a dark band corresponding to the position of the light which has been intercepted by the analyser.

In the case of a solution of an active substance in an inactive solvent, the rotation is proportional to the quantity of the substance present in the solution. If a is the rotation produced at any given temperature for light of a given colour by a length l, *expressed in decimetres*, of a solution containing x grams of the active substance in one cubic centimetre of the solution, the quantity a/lx is called the specific rotation of the substance at the given temperature and for the given-coloured light.

412. Connection between Optical Activity and Chemical and Physical Nature.—We have mentioned in the previous section that some samples of quartz rotate the plane of polarisation to the right and some to the left. It is found that this difference in their optical behaviour, exhibited by different specimens of quartz, is connected with a difference in their crystalline form. Thus the ordinary form of a quartz crystal is a six-sided prism topped by a six-sided pyramid. The alternate solid angles where two pyramid faces meet two faces of the prism are, however, often bevelled off by small secondary faces or facets which are inclined to the main faces. In any given crystal, these facets all appear to slope towards the right or towards the left when the crystal is held with the pyramid uppermost. When they slope towards the right, the specimen rotates the plane of polarisation to the right, and *vice versa*. A similar result was obtained by Pasteur with reference to the double racemate of sodium and ammonium.

Of amorphous bodies which exhibit optical activity, with the exception of one or two very little-known compounds containing nitrogen, it is found that they are not only all compounds of carbon, but that in every case one or more of the carbon atoms has its valency satisfied by four *different* atoms or radicals, which fact is generally expressed by saying that these bodies contain one or more asymmetrical carbon atoms.

There also exists in every case a twin substance, or isomer, which has the same composition but which rotates the plane of polarisation in the opposite direction. If the substance contains only one asymmetrical carbon atom, the rotation produced by the isomers are equal and opposite. Thus in the case of tartaric acid there is a dextro-tartaric acid which rotates the plane of polarisation to the right, and a levo-tartaric acid which rotates it to the left, and finally, in many reactions an inactive tartaric acid is produced. Pasteur has, however, shown by certain processes which only affect the dextro acid and not the levo, that this inactive acid consists of an equi-molecular mixture of the dextro and levo acids.

413. Use of Optical Rotation to Estimate Sugar—Saccharimetry.—Cane sugar being an optically active substance, if we measure the rotation produced by a known length of a solution of this substance, we can calculate from the specific rotation the quantity of sugar contained in the solution.

If the solution contains not only the dextro-rotatory cane sugar but also the levo-rotary levelose, after determining the rotation produced by the mixture, the cane sugar is converted into levelose by acting on the solution by means of hydrochloric acid, and then the rotation is again determined. The change in the rotation will give the quantity of cane sugar in the original solution.

BOOK V

MAGNETISM AND ELECTRICITY

PART I—MAGNETISM

CHAPTER I

MAGNETS AND MAGNETIC FIELDS

414. The Loadstone.—It was known to the ancients that certain ores of iron possess the property of attracting to themselves and retaining small particles of iron. This property was exhibited in a marked degree by some of the ores which came from a place in Asia Minor called Magnesia, and hence the ores which exhibited this property were called *magnetic stones.* All the phenomena connected with the properties of such magnetic stones, or magnets as they are now called, are referred to as magnetic phenomena, and the branch of physics dealing with this subject is called magnetism. The loadstone consists of equivalent proportions of the two oxides of iron, FeO, Fe_2O_3.

If a natural magnet, as a loadstone is often called, be dipped in iron filings, the filings will be found to adhere to the magnet in very characteristic tufts; these tufts are not uniformly distributed over the surface, but are much more marked at some parts of the surface, chiefly projecting corners, than at others.

415. Artificial Magnets.—In addition to the loadstone, there are other bodies which exhibit magnetic properties; chief among these are bars of hard steel which have been treated in a manner which we shall consider in detail in a subsequent section. Such a bar of steel is called an artificial magnet, but since we shall be dealing with artificial magnets exclusively we shall in future term it a magnet, and when it is dipped in iron filings it attracts them and forms tufts, but these tufts are almost exclusively confined to the two ends of the bar. The ends of the magnet where the power of attracting iron filings seems to be situated are called the poles of the magnet.

Another fundamental property of a magnet which also was known, at any rate to the Chinese, long ago is that when a magnet is suspended or

pivoted so that it can turn freely about a vertical axis, it will set itself in a definite direction, which is very nearly parallel to the meridian, that is, it sets itself in the north and south direction. It is found that it is always the same end of any given magnet that points towards the north pole, and hence this pole of the magnet is called the north-seeking pole, or simply the north pole, while the other pole is called the south-seeking pole, or south pole. The fact that we are able in this way to distinguish the two poles of a magnet shows that there must be some difference between the two poles.

416. Magnetic Attraction and Repulsion.—If a magnet is suspended by a fine thread, or pivoted on a point, so that it can turn freely in a horizontal plane, then, as we have already said, it will set itself in a direction which, in the absence of any disturbing force due to other magnetic causes, will be very nearly due north and south. If under these circumstances the north pole of another magnet is brought near the north pole of the suspended magnet, this latter will be repelled. If, however, the south pole of the magnet is brought near the north pole of the suspended magnet, this pole will be attracted. In this way we may verify the following general law : Two poles of similar name repel one another, while two poles of different name attract one another.

This gives us a ready means of ascertaining which pole of a magnet is the north pole ; for we have only to bring one of the poles near the north pole of a suspended or pivoted magnet, such as a compass-needle, when, if the north pole of the compass-needle is repelled, we know that the pole of the other magnet must be a north pole.

417. Permanent and Temporary Magnetism.—If a bar of soft iron is dipped into iron filings, or is suspended so as to be able to rotate, it will neither attract the filings nor will it set itself in the north and south direction, in fact it is not a magnet. If, however, a magnet is brought near one end of the bar of soft iron and the other end is then dipped in iron filings, it will now attract the filings, forming tufts in the same way as a magnet does. If the magnet is now removed, the iron at once loses its power of attracting the filings. We thus see that, in addition to the permanent magnetism exhibited by a steel magnet, we have temporary magnetism induced in a bar of soft iron when it is in the neighbourhood of a magnet. Other substances besides soft iron possess the property of acquiring temporary magnetism, though to a much smaller degree than in the case of iron. Such a body is called a *magnetic body*, while a body, such as a piece of hard steel, which is permanently magnetised is called a *magnet*.

If an unmagnetised piece of steel is brought near a magnet it will become magnetised, as did the piece of soft iron under the same conditions ; on the removal of the magnet the steel will, however, not lose its magnetism but will have become a permanent magnet. This difference between the behaviour of steel and soft iron is referred to as being due to

the superior coercive power of the steel. We shall return to this subject in a future section.

Magnets are made in many different shapes, according to the purpose for which they are intended. The two commonest shapes are a straight bar of which the section is either a rectangle or a circle, such a magnet being called a bar magnet, while the other form would be derived from a bar magnet by bending it round in the form of a horse-shoe, so that the north and south poles are brought near together; such a magnet is called a horse-shoe magnet. Other special forms of magnets we shall consider when dealing with the instruments in which they are used.

418. Magnetic Lines of Force.—We have seen that a magnet is capable of exerting a force on another magnet, even when they are separated by some distance, so that the space surrounding a magnet possesses some properties, due to the presence of the magnet, which it does not possess when the magnet is not present. We therefore say that a magnet is surrounded by a magnetic field of force, for in the space considered magnetic forces are brought into play.

If a small compass-needle is brought within the field of force of a magnet it will set itself at each point in a definite direction. If lines are drawn so that they are everywhere in the direction in which a compass-needle would set itself under the influence of the magnet, these curves are called the lines of force of the field of force of the magnet, or, more shortly, the lines of force of the magnet.

When a small pivoted compass-needle is placed in the neighbourhood of a magnet both its poles will be acted on by the two poles of the magnet. Thus the north pole of the needle will be attracted by the south pole of the magnet and repelled by the north pole, while the south pole of the needle will be attracted by the north pole of the magnet and repelled by the south pole. The needle will therefore set itself in such a direction that these four forces will have no resultant moment round the pivot about which the needle can turn. Hence the resultant of all four forces must have no moment round the pivot (§ 73).

Let NS (Fig. 397) be the magnet and *sn* the needle in its position of equilibrium under the influence of NS. Then the forces acting on the needle are along *s*A, *s*B, *n*C, *n*D. If the length of the needle is sufficiently small the two forces *s*A and *n*C are parallel and equal and opposite, for the points *n* and *s* may be taken as at the same distance from N. In the same way the two forces *s*B and *n*D are equal and opposite. Hence the resultant *s*R is parallel to the resultant *n*R′. Therefore if these two resultants are not to have a couple about the pivot of the needle K, they must both act in a direction passing through K, that is, the line joining the poles of the needle must be parallel to the resultant force acting on a magnetic pole placed at the point K. Hence the direction of the line of force at any point represents the direction of the resultant magnetic force due to the action of the two poles of the magnet.

The form of the lines of force can be most easily obtained by making use of the magnetism induced in iron filings when they are placed near the magnet, for each filing becomes magnetised and tends to set itself parallel to the lines of force. Hence if iron filings are scattered over a

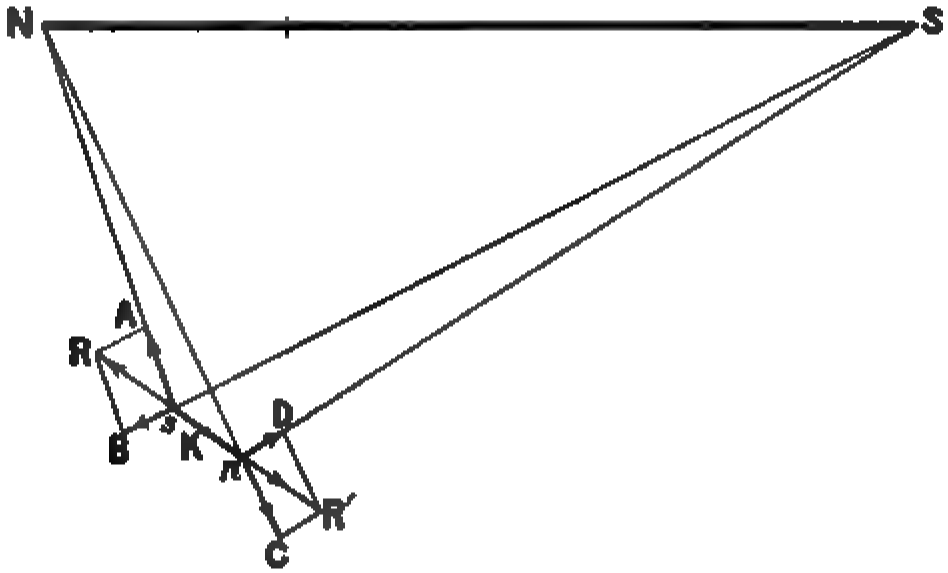


FIG. 397.

sheet of smooth paper or glass laid over the magnet, and then the paper or glass is lightly tapped, so as to facilitate the turning of the filings into the position in which their length is parallel to the lines of force, the form of these latter can at once be obtained. The curves shown in FIG. 398,

FIG. 398.

which represent the lines of force of a bar magnet, have been obtained in this way.

Since a force must have not only direction but also sense, we have to adopt some convention as to the sense in which the lines of force are taken to act, for while a north pole would be acted on by a force tending

to move it along a line of force in one direction, a south pole would be acted on by a force in the opposite direction. We shall take the direction in which the force would tend to move a north pole to be the positive

FIG. 399.

direction of the line of force, so that the lines of force will run from the north pole of the magnet to the south.

Each line of force will therefore start from a north pole and end at a south pole. If now the lines of force between a north pole of one magnet and the south pole of another are obtained it will be found they have the general form shown in Fig. 399. Here it will be seen that some of the

FIG. 400.

lines of force run from the north pole of one magnet to the south pole of the other. Hence if the lines of force were physical realities, and there existed a tension along each line, the two poles in Fig. 399 would be

drawn together by this tension. The attraction between two unlike poles can therefore be referred to the existence of a tension existing along the lines of force. Since in the case of a single magnet the lines of force do not all stretch straight from one pole to the other, but are spread out, we must further suppose that each line of force exerts a repellent force on the neighbouring lines of force.

In the case when the two poles are of the same kind, the lines of force have the form shown in Fig. 400, and the repulsive force exerted by the lines of force one on the other tends to force the two poles apart, thus producing the repulsion that takes place between two poles of the same kind.

419. Fields of Magnetic Force.—The region of space surrounding a magnet in which magnetic phenomena are exhibited is called a magnetic field, the lines of force showing the direction in which the magnetic forces act. The fact that a suspended or pivoted magnetic needle, even when no magnet is in the neighbourhood, sets itself in a definite direction, shows that the space on the surface of the earth must be a magnetic field, and we shall find later on that there are other ways of producing magnetic fields besides placing a magnet in the neighbourhood.

A magnetic field in which the lines of force are all parallel is called a uniform field of force. In the absence of magnets or of magnetic bodies the field due to the earth's magnetism is, over any moderate area, practically uniform.

If a magnet is placed in a uniform field, the lines of force of the field will be distorted by the lines of force of the magnet in such a way that they are everywhere parallel to the direction of the resultant magnetic force due to the original field and to the magnet.

FIG. 401.

In Fig. 401 the directions of the lines of force due to the disturbance of a uniform field by the magnet NS are shown, the arrow representing the direction of the lines of force in the undisturbed field. It will be observed how some of the lines of force due to the field enter the south pole of the magnet, while some of the lines of force which leave the north pole of the magnet do not return to the south pole, but pass off to replace the lines of the

external field which have been absorbed by the magnet. The tension of the lines of force which, while they begin or end in the magnet, pass off to the field, will tend to turn the magnet in the anti-clockwise direction. Also, the repulsion between the lines of force of the magnet and of the field in the regions such as *a*, *a'*, will tend to turn the magnet in the same direction. Hence it can be seen how it is that a magnet tends to set itself with the axis, that is, the line joining the two poles, parallel to the lines of force of the magnetic field in which it is placed. The turning effect produced by one magnet on another is also illustrated in Fig. 402. Some of the lines of force which leave the north pole of the magnet NS enter the south pole of the magnet *ns*, while some of those which leave the pole *n* enter the pole S, and the tension along these lines tends to turn the magnets into the position in which their axes are parallel and their poles are pointing in opposite directions. If, instead of a magnet, we place a cylinder of soft iron in a uniform magnetic field, the lines of force will crowd together, entering the iron at one end and leaving it at the other end in the manner shown in Fig. 403.

FIG. 402.

Since the point where the lines of force enter a magnetic body is a south pole, while the point where the lines leave it is a north pole, it is evident that the end *n* of the iron cylinder becomes, under the influence of the field, a north pole, while the end *s* becomes a south pole.

FIG. 403.

It would thus appear that the lines of force of the field prefer passing through the iron to passing through the surrounding air, for they crowd into the iron. This crowding of the lines of force into the iron is also illustrated in Fig. 404, which represents the lines of force of a uniform field disturbed by a hollow iron cylinder. It will be noticed that the

filings within the cylinder are not oriented in any definite directions. This indicates that none of the lines of force cross the air within the cylinder, so that the soft iron has shielded the space within from the effects of the magnetic field.

420. Molecular Magnets.—If a long thin bar magnet is tested by plunging it in iron filings, these will be found to attach themselves almost exclusively at the ends or poles. Also, if the directions of the lines of force for the magnet are drawn, it will be found that almost all the lines of force leave the magnet near one end and enter it near the other. If now the magnet is broken in two parts and each of these is again tested, it will be found that each is a perfect magnet, having a north and a south pole. Hence, although the part of the bar which was originally a north pole is one still, the other end of the half bar, which in the whole magnet did not exhibit the properties of a pole, is now a south pole, while the portion of the other half, which in the whole magnet was not a pole, is now a north pole.

FIG. 404.

If now the broken magnet is put together in the position which it occupied before it was broken and is again tested with filings, and the lines of force are drawn, it will be found that again the centre hardly

FIG. 405.

exhibits any signs of poles, that is, the filings will not adhere to any great extent at the centre, nor will the lines of force enter or leave the reunited magnet at the centre. The reason for this is evidently that the effects of the north pole which exists at one side of the break is neutra-

lised by the south pole which exists at the other side, all the lines of force which leave the north pole entering the neighbouring south pole, and none straying out into the surrounding air. If each of the halves is again broken into two, it will be found in the same way that they are each a complete magnet, with a north and a south pole. In Fig. 405 is given the iron-filing picture of the lines of force of a magnet which has been broken into four pieces, the pieces having been placed at a little distance from one another. Fig. 406 gives the corresponding picture in the case when the pieces are placed close together. Proceeding in this way, it is found that however small the subdivisions into which the magnet is broken, the parts are each a complete magnet, having a north and a south pole. Hence we are led to the idea of molecular magnets, that is, that the molecules of a substance such as steel are all small magnets. In the unmagnetised state we may suppose that the molecular magnets have their axes pointing in all directions, so that the north pole of one is neutralised by the south pole of one of the neighbouring molecules. Fig. 407 represents the filing figure for a number of small magnets arranged with their axes turned in all directions, and it will be observed that the filings surrounding the magnets are very little affected. In a magnetised bar of steel, however, a greater

Fig. 406.

Fig. 407.

proportion of the small magnets are turned with their north poles pointing in one direction, and the greater the proportion of the molecular magnets which are turned in this direction, the stronger is the magnetisation of the bar. This is illustrated in Fig. 408, which gives the filing picture for the same magnets as in Fig. 407, but now they are all arranged with their axes pointing in one direction.

FIG. 408.

The fact that in a magnet the magnetic force is restricted to near the ends or poles is easily explained on this hypothesis, for in the case of the molecular magnets in a magnetised bar it is evident that, except at the ends of the bar, the north pole of each small magnet will be very near the south pole of the next magnet, and hence these two will neutralise each other's effects on all external points. At the two ends, however, this neutralisation will not occur, and at one end the north poles will combine to form the north pole of the magnet, and at the other end the south poles will combine to form the south pole.

If further we suppose that, in the case of steel, the molecular magnets having been once set in one direction they will remain in this direction, while in the case of soft iron, although under the influence of a magnetising force, the molecular magnets can be turned so that they lie in one direction, yet when the magnetising force is removed, the molecular magnets do not remain in their regular arrangement, but again turn in all directions, the difference in the behaviour of steel and iron can at once be accounted for. We shall in a subsequent section return to this subject, and show how this hypothesis of molecular magnets is capable of explaining the magnetisation of iron, even when considered in greater detail.

421. Coulomb's Law.—Although, as we have seen when considering the effects of breaking a magnet in bits, we are unable to obtain either a north pole or a south pole alone, yet, if we have a very long magnet, in the space surrounding one of the poles the magnetic forces are practically due to that pole alone, for the other pole is at such a great distance that it produces practically no effect. Hence we can in this way get what is practically a single pole, and it is very convenient in considering the subject to speak of a single pole, and of the forces which act on such a single pole.

Coulomb examined the laws governing the attraction and repulsion between magnetic poles by suspending a long thin magnet by means of

a wire, the upper end being attached to a divided head, so that the angle through which the top of the wire was turned could be read off. A second long thin magnet was placed with its axis vertical, and one of its poles in the same horizontal plane as the suspended magnet. The force with which the pole of the fixed magnet repels or attracts one of the poles of the suspended magnet, when at different distances, was measured by finding the angle through which the torsion head had to be turned to keep the poles at the given distance apart (see §§ 109, 174).

From the results of this series of experiments, carried on by means of the torsion balance, Coulomb found that the force exerted between two poles was proportional to the product of the strengths of the poles, and inversely proportional to the square of the distance between the poles. We shall see later that the force also depends on the nature of the medium between the poles, if, however, we suppose the intervening medium always to be air, then the force F, exerted between two poles of strength m and m', is, according to Coulomb's law, given by the equation

$$F = \frac{mm'}{r^2} . k,$$

where r is the distance between the poles, and k is a constant.

422. The Unit Magnetic Pole.—Coulomb's law gives us a means not only of measuring the strength of magnetic poles, but also of defining the unit pole. If we take two poles of the same strength m, and place them at *unit* distance apart in air, then the force exerted between them will be given by $F = m^2k$. If further we choose m in such a way that the repulsion between the two poles is the unit of force, we have $m^2k = 1$. If we define our unit pole as such that when two unit poles are placed at a distance apart of one centimetre in air they repel each other with the force of one dyne, then k will be equal to 1, and Coulomb's law may be expressed symbolically by the equation

$$F = \frac{mm'}{r^2}.$$

423. Magnetic Moment.—Although it is convenient for the theoretical discussion of the subject to speak of a single magnetic pole, yet in practice such a thing does not exist, and as the forces in play between two magnetic poles can only be measured by determining the force acting on a magnet, which must of necessity possess both a north and a south pole, it is convenient to have some quantity which shall include the effect of the two poles. Such a quantity is the product of the strength of either pole of the magnet into the distance between the poles, and is called the *moment of the magnet.* The reason why this quantity is of importance will be at once apparent if we consider that since the two poles of a magnet always act in opposition, the greater will be the combined effect the greater the distance between them, for then the magnetic effect of one pole will be less neutralised by that due to the other.

424. Strength of a Magnetic Field.—Now that we have defined the unit magnetic pole, we are in a position to define the strength of the unit magnetic field. The strength of a magnetic field is the force with which a unit pole would be acted upon by the field when placed at the given point. Thus if the strength of a magnetic field at a given point is H, a unit north pole when placed at the point will be acted upon by a force of H dynes in the direction of the lines of force at the point. If the strength of the pole is m, then it will be acted on by a force of mH dynes in the direction of the lines of force.

425. Couple Acting on a Magnet in a Magnetic Field.—Suppose that a magnet NS (Fig. 409) is placed in a uniform field, that is, a field in which the lines of force are everywhere parallel, and of which the strength (H) is everywhere the same. If YOY′ is the direction of the field, the magnet will be acted upon by a couple tending to turn it round in the clockwise direction. If the strength of each pole of the magnet is m, the pole N will be acted upon by a force mH in the direction $\overrightarrow{\text{NH}}$. The moment of this force about the point O is $m\text{H}.\overline{\text{NL}}$, where NL is the perpendicular from N to the line YOY′. In the same way the force acting on the south pole, S, will be equal to a force of mH in the direction $\overrightarrow{\text{SH}'}$, and the turning moment of this force about the point O will be equal to $m\text{H}.\overline{\text{SK}}$. Since O is the centre of the magnet, $\overline{\text{ON}}$ is equal to $\overline{\text{OS}}$, and the angle NOL is equal to the angle SOK, and therefore $\overline{\text{NL}}$ is equal to $\overline{\text{KS}}$.

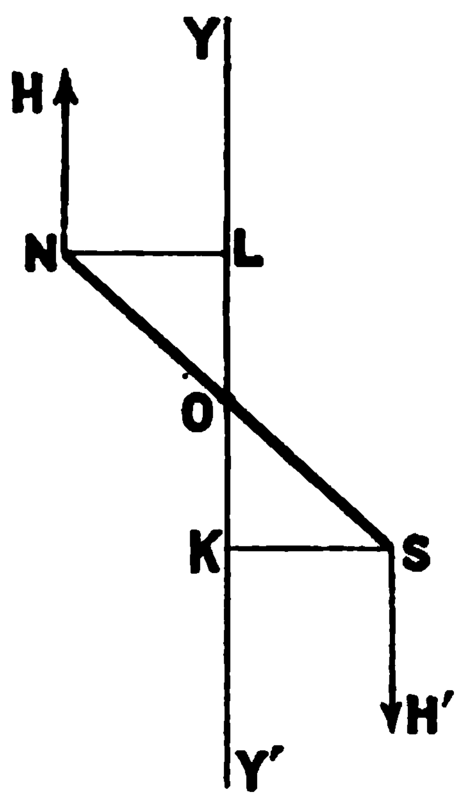


FIG. 409.

Since the forces $\overrightarrow{\text{NH}}$ and $\overrightarrow{\text{SH}'}$ are equal and opposite parallel forces, they constitute a couple (§ 70), and the turning moment of this couple is equal to the product of one of the forces into the perpendicular distance between the lines of action of the parallel forces. Hence the turning moment is equal to $mH(\overline{\text{NL}}+\overline{\text{KS}})$. Now if θ is the angle, NOL, between the axis of the magnet and the direction of the field, and l is the length of the magnet, so that $\overline{\text{NO}}$ is equal to $l/2$, we have

$$\overline{\text{NL}}=\overline{\text{ON}}\sin\theta=l/2.\sin\theta.$$

In the same way, $\overline{\text{KS}}=\overline{\text{OS}}\sin\theta=l/2.\sin\theta$. Hence $(\overline{\text{NL}}+\overline{\text{KS}})$ is equal to $l\sin\theta$. Thus the turning moment exerted by the field on the magnet is

$$mHl\sin\theta.$$

But the product of the strength of one pole of a magnet into the distance between the poles is the magnetic moment of the magnet. Hence, if M

is the magnetic moment of the magnet, $M = ml$, and the turning couple due to the action of the field on the magnet is $MH \sin \theta$.

The above expression will allow us to measure either M or H, if we know the other and can measure the couple acting on the magnet when it is turned so that its axis makes a known angle, θ, with the direction of the field.

Since the forces acting on the two poles of a magnet when it is placed in a uniform field constitute a couple, they have no resultant tending to produce a motion of translation in the magnet. This is proved experimentally by floating a magnet on a disc of cork, when the magnet turns and sets itself approximately north and south, but does not move off in any direction.

426. Couple due to the Action of one Magnet on another.—Let *ns* (Fig. 410) be a small magnet placed with its centre on the prolongation of the axis of another magnet, NS, so that the axes of the two magnets are at right angles to one another. Then the north pole N of

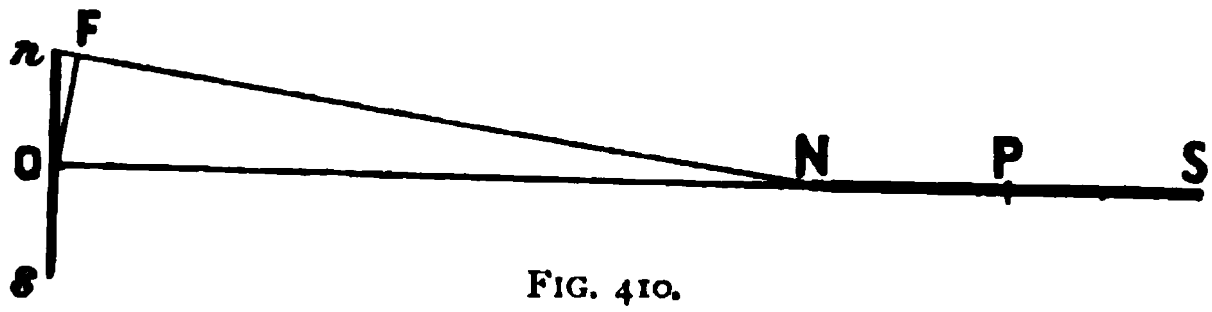


FIG. 410.

the one magnet will repel the north pole n of the other and attract the south pole s, while the south pole of the first magnet will attract n and repel s; but since the pole N is nearer than the pole S, the resultant action of the two will be a force acting on *ns*, tending to turn it in the anti-clockwise direction. Let m be the strength of one of the poles of NS, and m' the strength of one of the poles of *ns*, and let $2l$ be the length of *ns* and $2L$ that of NS. Then the repulsion between N and n is equal to $\frac{mm'}{r_1^2}$, where r_1 is the distance from N to n. The turning moment of this force about the point O will be equal to $\frac{mm'}{r_1^2} \times \overline{OF}$, where $\overline{OF}$ is drawn perpendicular to Nn. But if the distance, $\overline{OP}$, between the centres of the two magnets is called D, since the triangles nOF and nON are similar, $\frac{\overline{OF}}{\overline{On}} = \frac{\overline{ON}}{\overline{Nn}}$. Hence

$$\overline{OF} = \frac{(D-L)l}{r_1} = \frac{(D-L)l}{\sqrt{l^2 + (D-L)^2}}.$$

So that the turning moment is

$$\frac{mm'l(D-L)}{\{l^2 + (D-L)^2\}^{\frac{3}{2}}}.$$

The turning moment due to the action of N on s will be the same as that due to N on n. Thus the total turning moment due to the pole N will be

$$\frac{2mm'l(D-L)}{\{l^2+(D-L)^2\}^{\frac{3}{2}}}.$$

In the same way the turning moment due to the pole S will be obtained by writing $D+L$ for $D-L$ in the above expression, and since this moment acts in the opposite direction to that due to the pole N, the total turning moment due to the two poles will be

$$\frac{2mm'l(D-L)}{\{l^2+(D-L)^2\}^{\frac{3}{2}}}-\frac{2mm'l(D+L)}{\{l^2+(D+L)^2\}^{\frac{3}{2}}}.$$

If the length of the magnet ns is so small, compared with the distance D, that the term l^2 can be neglected in comparison with $(D-L)^2$, the above expression reduces to

$$\frac{2mm'l}{(D-L)^2}-\frac{2mm'l}{(D+L)^2},$$

or

$$\frac{2mm'l\{(D+L)^2-(D-L)^2\}}{(D^2-L^2)^2},$$

or

$$\frac{4mm'lDL}{(D^2-L^2)^2}.$$

If, further, the distance between the two magnets is so great compared with the length of either that we may also neglect L^2 compared to D^2, the expression for the turning moment acting on the needle ns reduces to

$$\frac{4mm'lL}{D^3}.$$

Now $2mL$ is the magnetic moment, M, of the magnet NS, while $2m'l$ is the magnetic moment, M', of the magnet ns. Hence the turning moment may be written

$$\frac{2MM'}{D^3}.$$

Next, suppose that the needle ns, instead of being placed with its axis perpendicular to the axis of NS, is placed so that its axis makes an angle

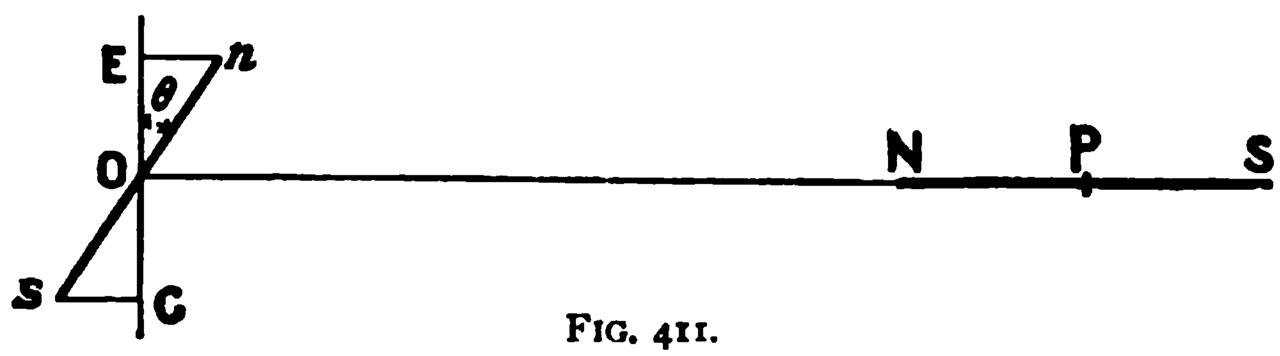


FIG. 411.

θ with the line drawn through its centre perpendicular to the axis of NS (Fig. 411). If we suppose that the magnets are both small compared to the distance between them, the turning moment will be less than before

in the ratio of 1 to cos θ. For we have now practically to do with a needle of which the strength of the poles, m', is the same as before, but the distance between the poles is now equal to $\overline{\text{EG}}$, where E and G are the feet of the perpendiculars drawn from n and s on the line which passes through O, and is perpendicular to the axis of the magnet NS. But $\overline{\text{OE}}=\overline{\text{O}n}\,.\cos\theta$, hence the moment of this imaginary magnet is $M'\cos\theta$. Thus the turning moment due to NS is now equal to

$$\frac{2MM'\cos\theta}{D^3}.$$

This position of two magnets, in which the axis of one is at right angles, or at any rate very nearly at right angles to the axis of the other, is called the "A tangential position of Gauss."

The above formula shows that when θ is a right angle, so that the axes of the two magnets lie in the same straight line, cos θ being zero, there will be no turning couple due to the action of one magnet on the other; this is also evident from a consideration of the direction of the lines of force of the magnet NS, which at O are parallel to the axis of NS.

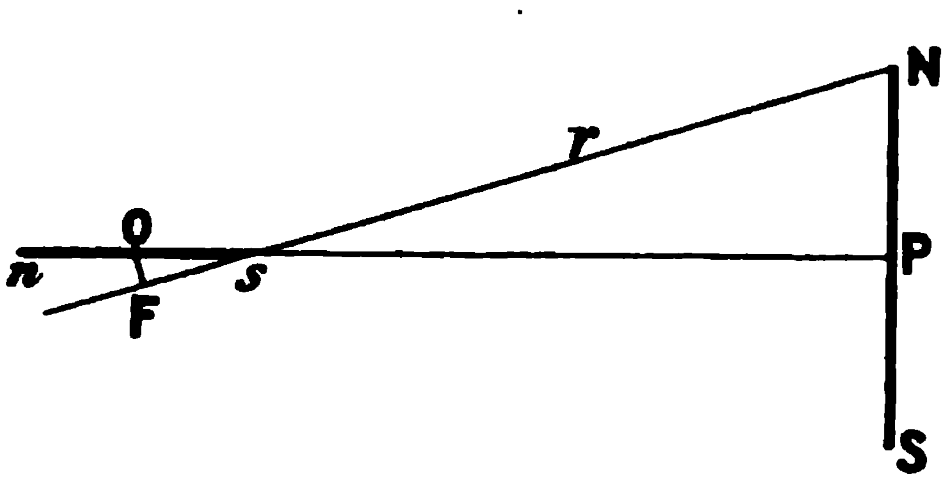


FIG. 412.

We have now to consider the case when the magnet, NS, is placed with its centre on the prolongation of the axis of ns and at right angles to this line as shown in Fig. 412. This is called the "B tangential position of Gauss." Using the same notation as before, the force exerted by N on s is equal to $\frac{mm'}{r^2}$ along $\overrightarrow{s\text{N}}$. The turning moment of this force about O is $\frac{mm'}{r^2}\times\overline{\text{OF}}$. Since the triangles OF$s$ and NPs are similar, $\frac{\overline{\text{OF}}}{\overline{\text{O}s}}=\frac{\overline{\text{NP}}}{\overline{\text{N}s}}$; or $\overline{\text{OF}}=\frac{lL}{r}$. But

$$r^2=\overline{NP}^2+\overline{Ps}^2=\overline{NP}^2+(OP-Os)^2$$
$$=L^2+(D-l)^2.$$

Hence the turning moment due to the action of N on s is

$$\frac{mm'lL}{\{L^2+(D-l)^2\}^{\frac{3}{2}}}.$$

This will also express the turning moment due to the action of S on s,

while the moments due to the action of N and S on n are obtained by changing $D-l$ into $D+l$ in this expression. In this case, however, it will be at once seen that the turning moments due to the action of the poles N and S on the two poles n and s of the needle are in the same direction. Hence the total turning moment is

$$\frac{2mm'lL}{\{L^2+(D-l)^2\}^{\frac{3}{2}}}+\frac{2mm'lL}{\{L^2+(D+l)^2\}^{\frac{3}{2}}},$$

or

$$\frac{MM'}{2\{L^2+(D-l)^2\}^{\frac{3}{2}}}+\frac{MM'}{2\{L^2+(D+l)\}^{\frac{3}{2}}}.$$

If now we proceed to take the needle short in comparison to the distance, D, between the magnets, this expression reduces to

$$\frac{MM'}{(L^2+D^2)^{\frac{3}{2}}}.$$

While if L^2 may also be neglected compared with D^2, the turning moment reduces to

$$\frac{MM'}{D^3}.$$

As before, if the axis of the needle, ns, makes an angle θ with the line $\overline{OP}$, the turning couple will be

$$\frac{MM'}{D^3}\cos\theta.$$

Again, this is zero when $\theta=90°$, *i.e.* when the two magnets are parallel, and as the lines of force of the magnet NS at the point O are parallel to the axis of NS, this is as it ought to be.

These expressions may be used to test the correctness of Coulomb's law, which we have employed in obtaining them, for if the needle ns is suspended by a wire or by a bifilar suspension (§ 119), then by turning the upper end of the wire, or of the bifilars, till the axis of the magnet ns comes into the position considered, we can, as has been explained in § 109, measure the couple which is acting on ns when the magnets are at a given distance, D, apart. Then by varying D we can test whether the couple varies inversely as D^3. We may also see if, for a given value of D, the couple, when the magnets are in the A position, is twice as great as the couple when they are in the B position.

Also by using two magnets, of which the moments are M and M″, we may measure the couple they exert when used separately at a distance D. We may then use them both simultaneously, placed at the distance D, one on either side of the needle, and show that the couples produced when they act so as to oppose each other, and also when they act to assist each other, and the couples which they each produce separately are in the ratio of $M-M'' : M+M'' : M : M''$.

This method of testing the truth of Coulomb's law can be made more accurate than can the method with the torsion balance, where the effect of the one pole is often modified by that of the other.

427*. Time of Vibration of a Magnet when Suspended in a Magnetic Field.—We have seen in § 425 that if a magnet NS (Fig. 413) of which the moment is M, is suspended in a magnetic field of strength H, and if the axis of the magnet makes an angle θ with the direction $\overrightarrow{AB}$ of the field, the couple acting on the magnet, and tending to turn it into the direction $\overrightarrow{AB}$, is $MH \sin \theta$. Thus if the magnet lies with its axis parallel to the direction of the field, and with its north pole pointing in the positive direction, it will be in stable equilibrium, and when it is displaced from this position through an angle θ, the couple tending to restore it to its undisturbed position will be $MH \sin \theta$. Hence if the magnet is displaced and then set free, it will, under the influence of this couple, move back towards its position of equilibrium. It will, during this motion, gain kinetic energy, so that it will pass through its equilibrium position and be displaced on the other side, and so on; in fact, it will execute oscillations about its position of equilibrium.

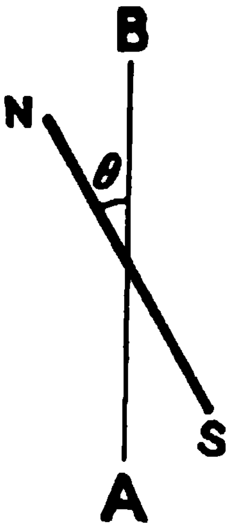


FIG. 413.

Now in § 113 we found that when a simple pendulum is displaced from its position of rest, the force tending to bring it back to that position was given by $mg \sin \theta$, where m is the mass of the bob. The similarity between this expression and that obtained in the case of the magnet is obvious, and we can at once see that if the simple pendulum performs isochronous vibrations, the suspended magnet will do so also. As we have already seen, the pendulum only performs isochronous oscillations when the maximum displacement (θ) is small, so that we may at once infer that the same will be the case with the suspended magnet.

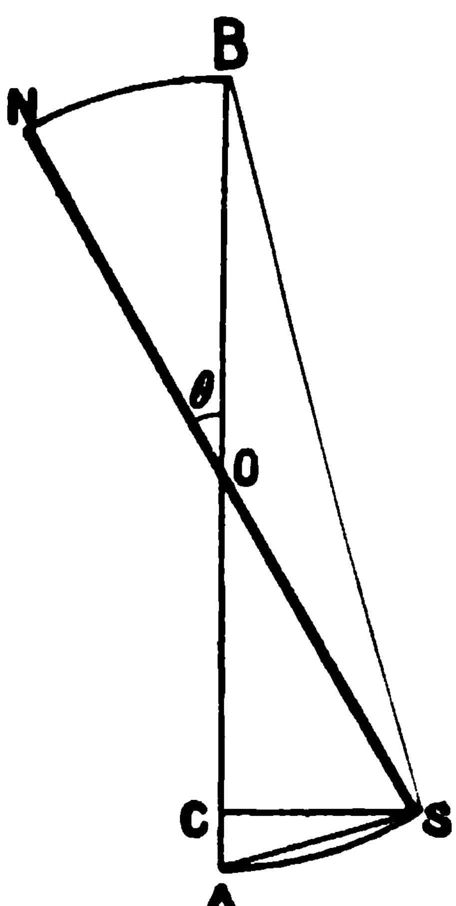


FIG. 414.

In order to find an expression for the periodic time T of the oscillations performed by the magnet, let AB (Fig. 414) represent the position of equilibrium of the magnet, and NS the position of the magnet when at its maximum elongation.

Now for small displacements (θ) the restoring force will be equal to $MH\theta$, since, as was shown in § 14 for small values of θ, $\sin \theta$ may be taken as equal to θ. Hence, for vibrations of small amplitude, the restoring force is proportional to

the displacement, and therefore the magnet will execute a simple harmonic motion (§ 50). Now in § 51 we have shown that in a S.H.M. the maximum linear velocity, that is, the velocity when the body is passing through its position of rest, is equal to $2\pi a/T$, where a is the amplitude of the vibration and T is the period. Hence the kinetic energy of a particle of mass m', when passing through its position of rest, is $2\pi^2a^2m'/T^2$.

Now the amplitude (a') of a particle at a distance r from the point O, about which the magnet rotates, is given by

$$\frac{a'}{(AS)}=\frac{r}{\overline{OA}}=\frac{2r}{l},$$

where (AS) stands for the *arc* AS, and l is the length of the magnet.

Therefore the kinetic energy of this particle, when passing through its position of rest, is

$$\frac{8\pi^2(AS)^2r^2m'}{T^2l^2}.$$

Thus the total kinetic energy of all the particles which build up the magnet is

$$\frac{8\pi^2(AS)^2}{T^2l^2}\Sigma(m'r^2)=\frac{8\pi^2(AS)^2K}{T^2l^2},$$

where K is the moment of inertia (§ 85) of the magnet about an axis through O.

Now when the magnet is at its extreme elongation the energy is entirely potential. This potential energy is equal to the work which has to be done to move the poles N and S into their new positions against the action of the field. Considering the pole S, we might take it from A to S along the paths AC, CS, where SC is perpendicular to AB, and therefore also perpendicular to the lines of force of the field. During the portion CS of the path, since the direction of motion is perpendicular to the force, no work is done. During the passage from A to C, since the force acting on the pole is mH, where m is the strength of the pole, the work done is $mH.\overline{AC}$. Hence the potential energy due to both poles when the magnet is at NS is

$$2mH.\overline{AC}.$$

Now the triangles ACS and ABS are similar. Therefore

$$\frac{\overline{AC}}{\overline{AS}}=\frac{\overline{AS}}{\overline{AB}}.$$

Thus the potential energy is

$$\frac{2mH.\overline{AS}^2}{l}.$$

Equating the potential energy at the extreme elongation, when the kinetic energy is zero, to the kinetic energy when the magnet is passing

through its position of rest, and therefore its potential energy is zero, we get

$$\frac{2mH\overline{AS}^2}{l} = \frac{8\pi^2(AS)^2}{T^2l^2}.$$

Now if the amplitude, θ, of the vibrations is small, the chord $\overline{AS}$ may be taken as equal to the arc (AS). Then

$$T^2 = \frac{4\pi^2K}{ml \, . \, H}.$$

But $ml = M$. Hence

$$T^2 = \frac{4\pi^2K}{MH},$$

or

$$T = 2\pi\sqrt{\frac{K}{MH}}.$$

428. Measurement of the Strength of a Magnetic Field.—We have seen in the last section that if a magnet, of which the moment is M and the moment of inertia is K, vibrates in a magnetic field of strength H, the periodic time T of the vibrations is given by

$$T = 2\pi\sqrt{\frac{K}{MH}}.$$

Hence if we measure T, and know K and M, we can calculate H. The moment of inertia K can either be calculated, if the magnet is of a simple and regular shape, or it can be determined experimentally. Hence we have only M and H to determine, so that if by any other experiment we

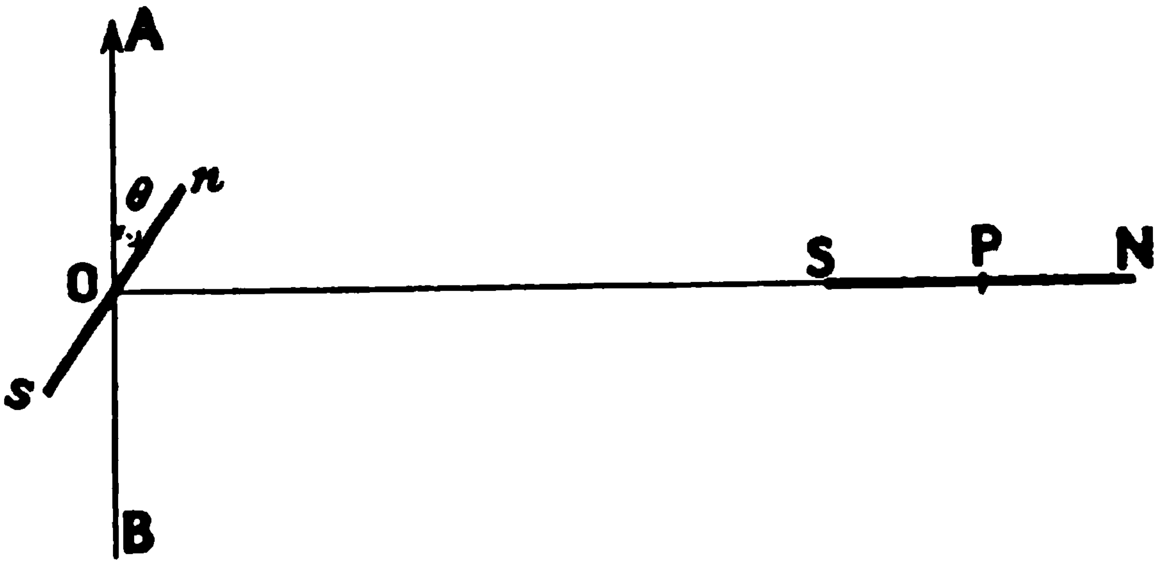


FIG. 415.

can get a second relation between M and H, say their ratio, we could calculate both of them.

Now we have obtained in § 426 expressions for the couple caused by the action of one magnet on another when they are placed in certain relative positions. Suppose now a magnetised needle ns (Fig. 415) is suspended by a fine thread in the given magnetic field, then it will set

itself parallel to the direction, $\overrightarrow{BA}$, of the field. If now we place the magnet, of which we have observed the period of vibration, in the position NS, it will exert a couple on the needle, which, if the distance OP is great compared to the sizes of the magnets, is equal to $\frac{2MM'}{D^3}$, and hence the needle will be turned into some such position as that shown in the figure, and will finally come to rest when the deflecting couple, due to NS, is equal to the couple, tending to bring it back into the direction $\overrightarrow{BA}$, due to the field.

If θ is the angle which the axis of the needle makes with the lines of force of the field when it comes to rest under the combined influence of the magnet NS and of the field, the couple acting in the clockwise direction due to the magnet is $\frac{2MM'}{D^3}\cos\theta$, while the couple acting in the opposite direction due to the field is $M'H\sin\theta$. When there is equilibrium these must be equal, and hence

$$\frac{2MM'}{D^3}\cos\theta = M'H\sin\theta,$$

or

$$\frac{M}{H} = \frac{D^3\sin\theta}{2\cos\theta} = \frac{D^3}{2}\tan\theta.$$

If then we measure the distance between the centres of the magnet and needle and the deflection, we can calculate the ratio $\frac{M}{H}$. But we have already seen that the vibration experiment gives us the value of the product MH, and hence by simple algebra the values of the two quantities M and H can be calculated. Therefore by measuring the periodic time of a magnet of known moment of inertia, when suspended in a given magnetic field, and then determining the angle through which a needle, suspended in the same field, is deflected by this magnet when placed at a known distance, we can obtain both the strength of the field and the magnetic moment of the magnet. Of course, when performing the deflection experiment, the magnet NS might be placed in the "B position," in which case $\frac{M}{H} = D^3\tan\theta$.

CHAPTER II

TERRESTRIAL MAGNETISM

429. The Magnetic Elements.—The most important magnetic field with which we have to do is that due to the magnetic state of the earth. In order to be able to state the condition of the magnetic field of the earth, or as we may say for short the earth's field, at any point we require to know two things, (1) the direction of the lines of force of the field, and (2) the strength of the field. That is, we want the direction in which a single unit north pole would tend to move under the influence of the field, and also the force which would act upon it. We have hitherto supposed that the directions of the lines of force of the magnetic fields with which we have been dealing were horizontal, so that a magnetised needle, which was suspended or pivoted, so as to turn about a vertical axis, was able to set itself parallel to the lines of force of the field. If a long thin unmagnetised bar of steel is suspended by a fine thread so that it hangs in a horizontal position, and is then magnetised, it will set itself in an approximately north and south position, but will no longer be horizontal. In this part of the globe the north end will dip downwards. This indicates that in these parts the lines of force of the earth's field are not horizontal, but are inclined downwards.

For most purposes it is convenient to suppose the earth's field resolved into two components, one of which is horizontal and the other vertical. Since a magnetic field is of the nature of a force, having magnitude or strength and direction, the field may be resolved into two component fields, just as a force in § 67 is resolved into two component forces.

In order to define each of these components, we require of course to know its direction and its strength. In the case of the horizontal component its strength is called the horizontal force, and is generally indicated by the letter H. Since by supposition this component is horizontal, in order to define its direction we only require to know the angle which it makes with some fixed direction. The fixed direction chosen is the geographical meridian, and the angle which the horizontal force makes with the geographical meridian is called the *declination*, or sometimes the *variation*.

The vertical component of the earth's field is called the vertical force, and is generally indicated by the letter V; its direction is along the vertical, *i.e.* the radius of the earth at the point considered.

The actual strength of the earth's field, which is of course the resultant of H and V, is called the total force. The angle between the lines of force of the earth's field and the horizontal is called the *dip*. Hence the dip is also the angle between the direction of the horizontal component and that of the total force or actual field. The three magnetic forces, the total force and its two components, H and V, must of course lie in the same vertical plane, the angle which this plane makes with a vertical plane containing the place considered and the axis about which the earth turns, that is, the meridian plane, is equal to the declination.

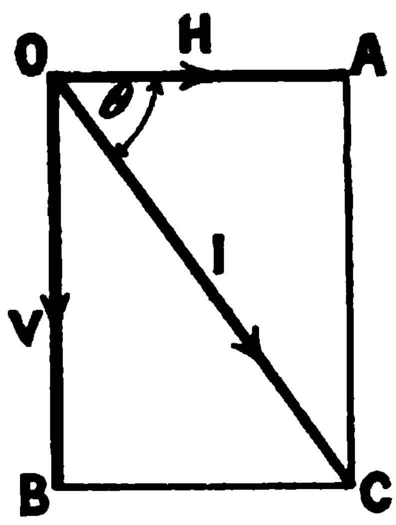


FIG. 416.

If the plane of the paper is taken as the vertical plane in which the total force and its components lie, and $\overrightarrow{OA}$, $\overrightarrow{OB}$, and $\overrightarrow{OC}$ (Fig. 416) represent in magnitude and direction the horizontal and vertical components and the total force, then the angle AOC or θ will be the dip. Hence if the total force, $\overrightarrow{OC}$, is called I, we have from the triangle AOC—

$$\frac{H}{I}=\cos\theta.$$

Also from the triangle BOC, since the angle BOC is $90^\circ-\theta$,

$$\frac{V}{I}=\cos \mathrm{BOC}=\sin\theta,$$

and finally

$$\frac{V}{H}=\tan\theta.$$

These three expressions permit of our obtaining V and I if we know the horizontal component, H, and the dip, θ, or if we know V and H we can obtain I and θ.

Hence it is evident that if we know the declination, the horizontal component, and the dip, we can deduce the direction and strength of the earth's field. Since it is generally most convenient to measure these three quantities, they are called the *magnetic elements*.

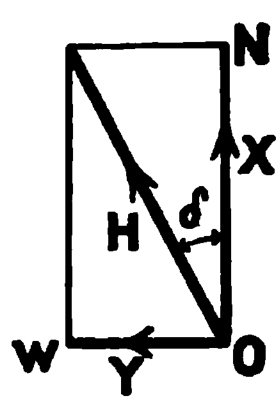


FIG. 417.

It is sometimes convenient to be able to express the direction and magnitude of the earth's field by three quantities which are all of the same nature, and not, as we have done above, by means of a force and two angles. Suppose we resolve the horizontal component along a line which points to the true or geographical north, and along a line true west. If X is the northerly component and Y the

westerly component, then it is at once evident, from Fig. 417, that if δ is the declination,

$$X = H\cos\delta,$$
$$Y = H\sin\delta,$$
$$\frac{Y}{X} = \tan\delta.$$

Hence, if we know X and Y, we can calculate H and δ. Thus the values of the three components of the force, X, Y, and V, are sufficient to completely define its value both in magnitude and direction.

Before proceeding to consider the general form of the earth's field as deduced from a study of the measurements which have been made of the magnetic elements at different parts of the earth's surface, it will be useful to briefly consider the methods employed to measure the magnetic elements at any given place.

430. Measurement of the Declination.—The declination is the angle between the geographical meridian and the direction of the horizontal component. Thus, since a magnet when suspended by a fine thread, so as to turn freely about a vertical axis, will set itself parallel to the direction of the lines of force of the horizontal component, the declination can be obtained by measuring the angle between the axis of such a suspended magnet and the meridian.

The practical difficulty in performing the experiment lies in the fact that the magnetic axis of a magnet does not necessarily coincide with its geometrical axis.

The magnet usually employed consists of a hollow steel cylinder, A (Fig. 418), which is fixed in a brass collar to which are attached two pegs, B and C, either of which fits into a clip attached to the end of a fine thread formed of unspun silk. At one end of the hollow magnet is placed a fine scale, S, engraved on a piece of glass, while at the other end is placed a lens, L. The focal length of this lens is equal to the length of the cylinder, so that the rays of light proceeding from any point in the scale S leave the lens as a parallel pencil. The line joining the central division of the scale and the optical centre (§ 348) of the lens is taken as the geometrical axis of the magnet.

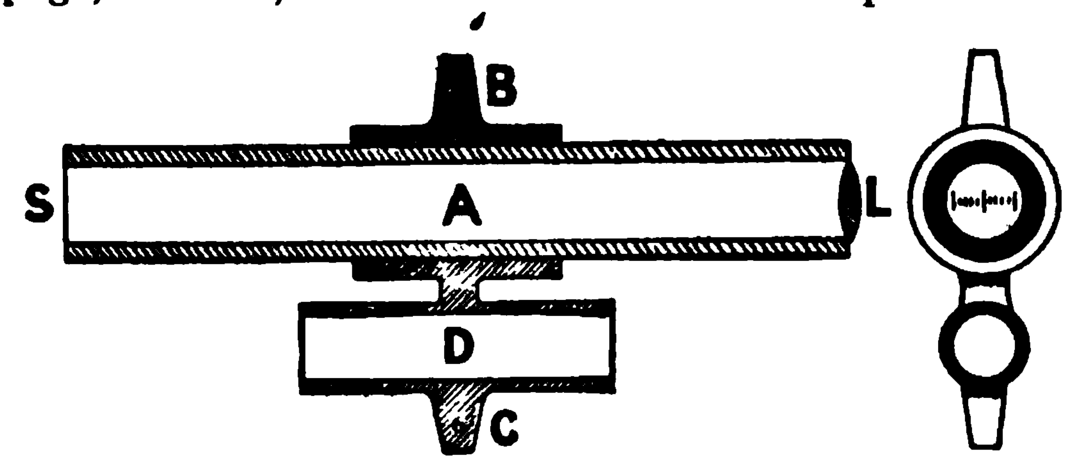


FIG. 418.

If AB (I., Fig. 419) is the plan of a magnetic needle suspended by a fine thread attached at C, and of which the magnetic axis is *ns*, then it will set itself with the magnetic axis in the magnetic meridian NS. In

this case the geometrical axis of the needle points to the west of the magnetic meridian. If the needle is reversed, so that what was the lower side is now the upper, then, as is shown at II., the geometrical axis AB will point as far to the east of the magnetic meridian as it did before to the west. Hence the magnetic meridian is half-way between the positions of the geometrical axis before and after the reversal of the magnet.

The cylindrical magnet shown in Fig. 418 is suspended in a box fixed to the centre of a divided circle, while an arm attached to the circle carries a small telescope, in the eye-piece of which are two intersecting

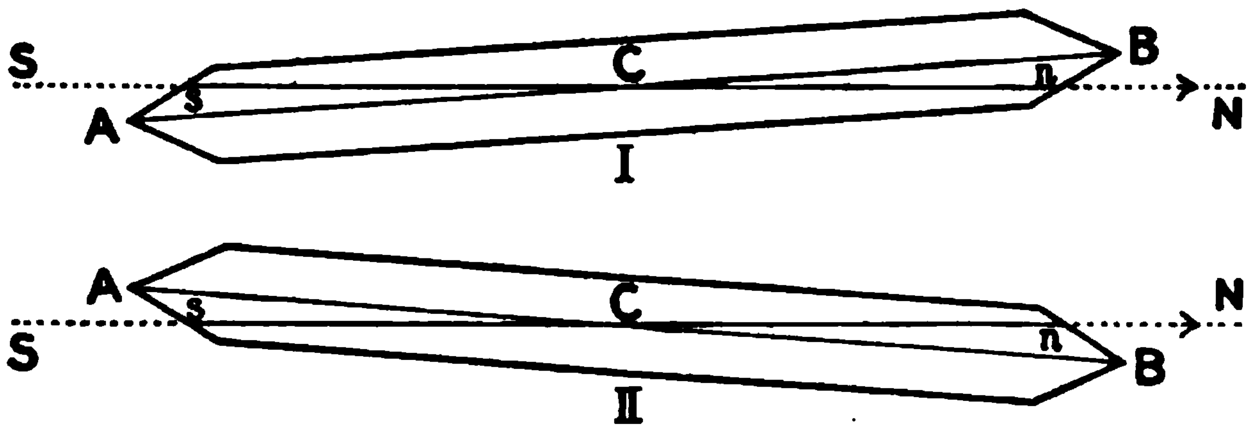


FIG. 419.
(From Watson's "Elementary Practical Physics.")

cross-wires. The telescope is turned till the centre division of the scale, S, coincides with the vertical cross-wire, first when the magnet is suspended by B, and then when it is reversed and is suspended by C. The mean of the readings on the divided circle corresponding to these two positions gives the reading corresponding to the magnetic axis of the magnet. The reading corresponding to the geographical meridian is obtained either by turning the telescope to view some object the bearing of which is known, or by observing the time of transit of a star or the sun over the vertical cross-wire

431. Determination of the Dip or Inclination.—When determining the declination by means of a magnet suspended by a thread, the effects of gravity on the magnet do not influence the observations, for the weight of the magnet will have no effect in producing a rotation about the thread as an axis.

In order to determine the dip, however, we have to support a magnet so that it can turn freely about a *horizonta'* axis, and then measure the angle which its magnetic axis makes with the horizontal. If the axis about which the magnet is allowed to turn passes through the centre of gravity of the magnet, the weight will have no moment round this axis, and will therefore not affect the position of the magnet. Since, however, it is practically impossible to secure this condition, the observations have to be so arranged that errors due to small departures from this condition may be eliminated. The principle is similar to that employed in the case of the determination of the declination, viz. to take readings in pairs,

such that the error in the separate readings affects the result in the opposite way, and hence the mean of the two readings gives the true value.

Suppose AEBD (Fig. 420) is the needle, the axle being at C, while the centre of gravity is at G, a point which does not coincide with C. We may consider the effect of this displacement of the centre of gravity as split up into two parts, one a displacement along the axis of the needle to K, and the other a displacement at right angles to the axis to H.

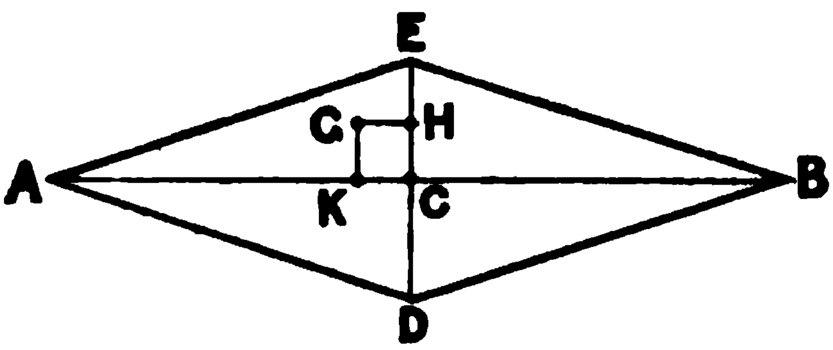


FIG. 420.

First consider the displacement of the centre of gravity at right angles to the axis of the needle, *i.e.* to H. If the end A is dipping, that is, in the northern hemisphere, if A is a north pole, and the needle is placed in its bearings, so that H is above C, the effect of the weight of the needle will be to increase the measured dip, while if the needle is reversed in its bearings, so that H is below C, the weight will decrease the measured dip by the same amount. Hence the mean of the readings obtained will be free from error, due to the displacement of the centre of gravity at right angles to the axis.

As long as the end A is dipping, the displacement of the centre of gravity along the axis to K will always increase the measured angle of dip. If, however, we remagnetise the needle, so that the end B dips, then the displacement of the centre of gravity to K will decrease the measured dip, so that by reversing the polarity of the needle this error can be eliminated. The fact that the magnetic axis of the needle may not coincide with the line joining AB is eliminated by the reversal of the needle when eliminating the effect of the displacement of the centre of gravity perpendicular to the axis, for the same reasons as in the case of the declination.

In order to measure the angle of dip, the needle is placed with its axle resting on two small horizontal agate knife-edges, K, K′ (Fig. 421), so that the axle is at the centre of a graduated circle, any slight want of agreement between the position of the axle and the centre of the circle being eliminated by reading the position of both ends of the needle by means of the two microscopes, M, M′. When not in use the needle is raised from the agates by means of two V-shaped supports, LL′, which can be raised by turning the knob E.

The only remaining source of error which may occur, owing to the imperfect adjustment of the instrument, is that due to the fact that the line joining the zero gradations on the circle may not be truly horizontal. The error due to this cause can be eliminated by taking two sets of readings with the instrument turned so that the graduated side of the circle

faces first east and then west. For the measured dip will in one case be greater than the true dip, and in the other case less by the angle which the line joining the zero makes with the horizontal.

The angle between the magnetic axis of the needle and the horizontal is only equal to the dip when the vertical plane in which the needle can turn coincides with the magnetic meridian, that is, when the axle of the needle points due magnetic east and west. In order to secure this condition, the circle is turned till the needle is vertical. The needle being

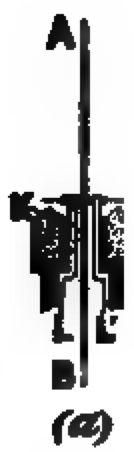


FIG. 421.

vertical shows that in the plane in which it can move there is no horizontal component of the earth's magnetic field. Now obviously this can only occur when the plane in which the needle moves is at right angles to the direction of the horizontal component, that is, at right angles to the magnetic meridian. We may also see that this must be so from the following reasoning : When the axle about which the needle turns is parallel to the horizontal component, this component can have no moment tending to produce rotation about the axle. Hence, when the plane in which

LINES OF EQUAL DECLINATION FOR THE YEAR 1900.

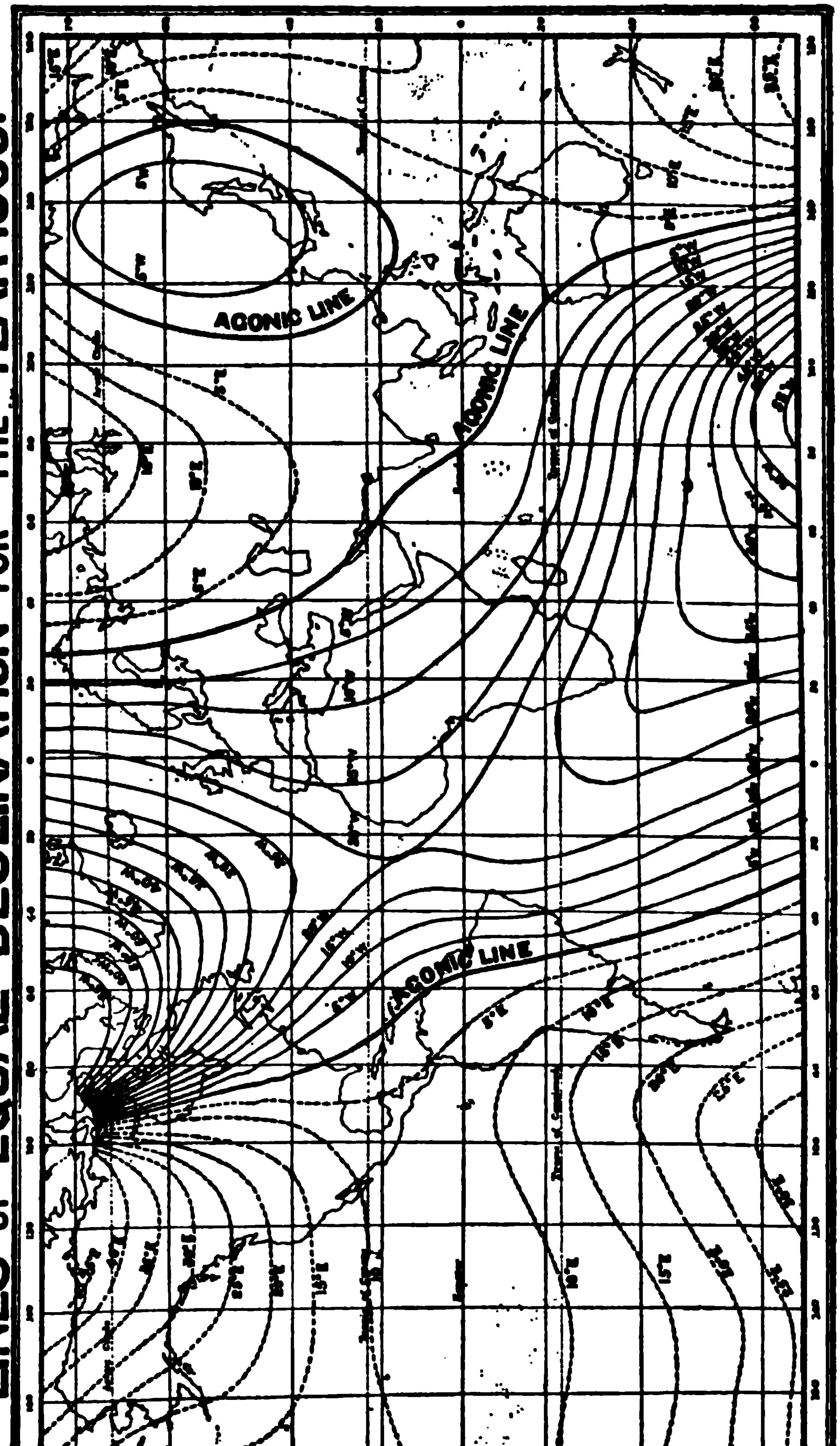


FIG. 422.

the needle turns is at right angles to the magnetic meridian, the vertical component is the only one which has a directive influence on the needle, which therefore sets itself in a vertical position. The position of the circle when the needle is vertical is read off on a horizontal divided circle, FG, attached to the stand, and then by means of this same horizontal circle the vertical circle, and with it the uprights carrying the needle, is turned through 90°, which brings its plane into the magnetic meridian.

432. Measurement of the Horizontal Force.—In order to measure the horizontal force, the usual method employed is that given in § 428. The magnet shown in Fig. 418 is first allowed to oscillate, and its period of vibration is determined; it is then used to deflect another suspended magnet, and the value of H is deducted from the results of these two experiments by the method given.

The moment of inertia of the magnet used in the vibration experiment is obtained by taking the period when a brass cylinder, the moment of inertia of which can be calculated, is placed in the tube D (Fig. 418), and also taking the period without this brass cylinder. From these two observations the moment of inertia of the magnet and its appendages can be calculated.

433. Terrestrial Magnetic Lines.—The magnetic state of the earth is best shown by constructing magnetic maps, in which lines are drawn through the places at which the element considered has the same value. In the case of declination, lines drawn so that the declination is the same at all places through which they pass are called *Isogonal Lines*, or Lines of equal Variation or Declination.

The form of the isogonals for the year 1900 is shown in Fig. 422, which represents the earth on Mercator's projection. The full lines indicate westerly declination, *i.e.* the north end of the needle points to the west of true north, while the dotted curves indicate easterly declination.

The thick lines, which separate the regions in which the declination is westerly from the regions in which it is easterly, are lines where the declination is zero, and therefore the compass-needle points due north. These lines are called the *agonic lines*.

There are two distinct agonic lines. One of these runs down the west side of North America, cuts off a part of South America, and then passes to the Antarctic Ocean; reappearing on the other side, it passes through the extreme west of Australia, through the Persian Gulf, near the Crimea, and finally enters the Arctic Ocean near the North Cape, and presumably joins the other branch in North America. The other agonic line forms an oval curve, the greater part of which lies in Siberia, and is known as the Siberian Oval.

The lines of equal dip are called *isoclinal lines*, and their form for the year 1900 is shown in Fig. 423.

It will be seen that the line of zero dip, or, as it is called, the magnetic equator, forms a circle which agrees approximately with the geographical

LINES OF EQUAL DIP FOR THE YEAR 1900.

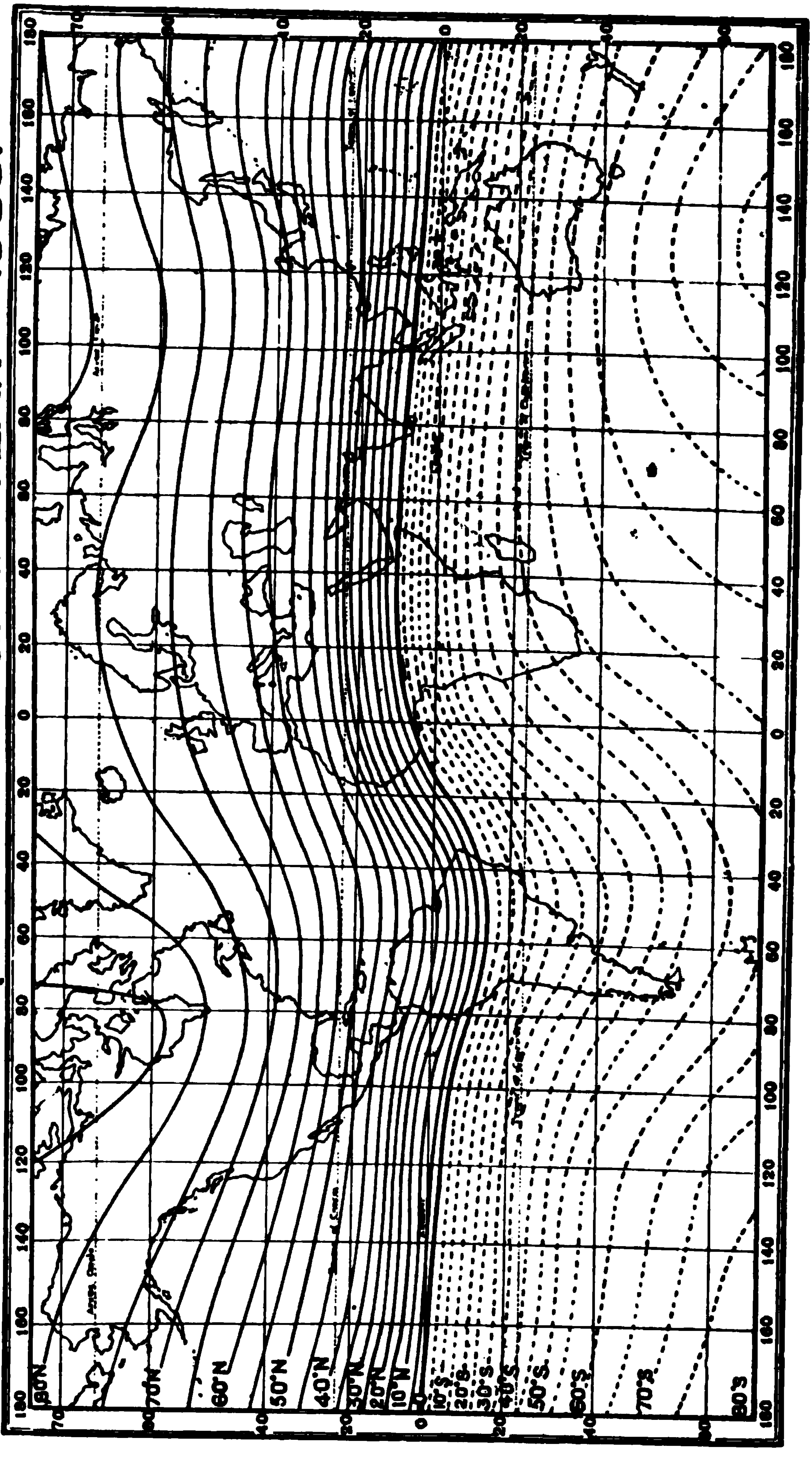


FIG. 423.

Fig. 424.

equator, and that the dip increases with the latitude, the north pole of the needle dipping in the northern hemisphere, and the south pole in the southern hemisphere.

At two points on the earth's surface the dip is 90°, *i.e.* the dipping-needle is vertical, this indicating that the horizontal component is zero, so that the compass-needle, which indicates the direction of the horizontal component, will not set itself at these points in any definite direction. These points are often called the magnetic poles. The north magnetic pole lies in lat. 70° 5′ N. and long. 96° 43′ W., while the south magnetic pole is at lat. 73° 30′ S. and long. 147° 30′ W.

The agonic line passes through the magnetic poles.

The lines of equal horizontal force are shown in Fig. 424, and it will be seen that the horizontal force is a maximum near the equator, and is zero at the magnetic poles.

The curves of equal total force are called *isodynamic lines*. The total force is not a maximum at the magnetic poles, but there exist in the northern hemisphere two points at which the total force is a maximum, while two similar points exist in the southern hemisphere. These points of maximum force are called magnetic foci. One of the northern foci is situated in North America at lat. 52° N., and long. 90° W. The other northern focus is at lat. 70° N. and long. 115° E., and is called the Siberian focus. The two southern foci are situated much nearer together than are the northern ones, their positions being approximately lat. 65° S., long. 140° E., and lat. 50° S., long. 130° E.

The positions of the terrestrial lines for the whole globe are necessarily only roughly known, for there are very large tracts where few, if any at all, determinations of the magnetic elements have been made. In the case of some more or less restricted portions of the earth, notably Great Britain, the magnetic elements have been determined with great accuracy at a large number of places, and hence the terrestrial lines are known with some accuracy. In Fig. 425 the lines of equal declination, dip, and horizontal force are given as obtained in an extensive magnetic survey conducted by Professors Rücker and Thorpe. These lines are obtained by combining the results of the measurements made at a number of stations which are grouped together so as to eliminate the effects of any local abnormality in the value of the elements at any one point, and hence they give what may be considered as the normal distribution of the lines. The value of the elements at any given spot do not, however, in general agree exactly with the values as deduced from these curves. This difference is due to the fact that at the place considered there may exist slight abnormalities, owing to the presence of magnetic material in the neighbouring portions of the earth's crust. The extent to which these disturbing causes may affect the even trend of the terrestrial lines is well shown in Fig. 426, which gives the form of the true isogonal lines, that is, the lines passing through the places at

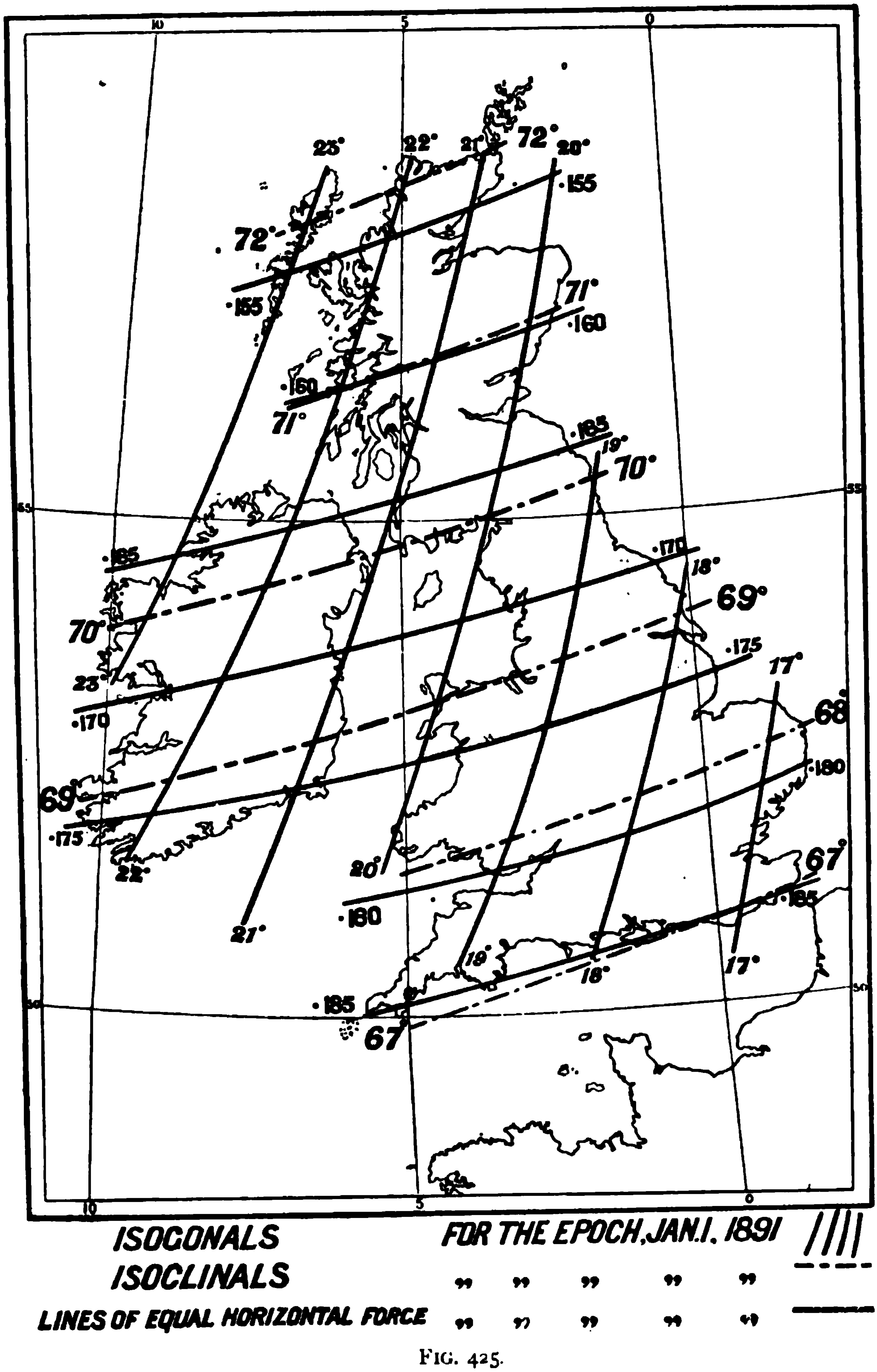

10
5
0
23°
22°
21°
72°
20°
·155
72°
·155
71°
·160
·160
71°
·165
19°
70°
·165
·170
18°
69°
70°
·175
17°
23°
·170
68°
·180
69°
·175
22°
67°
20°
·180
21°
·185
19°
18°
17°
·185
67
ISOGONALS FOR THE EPOCH, JAN. I. 1891
ISOCLINALS " " " " "
LINES OF EQUAL HORIZONTAL FORCE " " " " "

FIG. 425.

ISOGONALS FOR 1891.

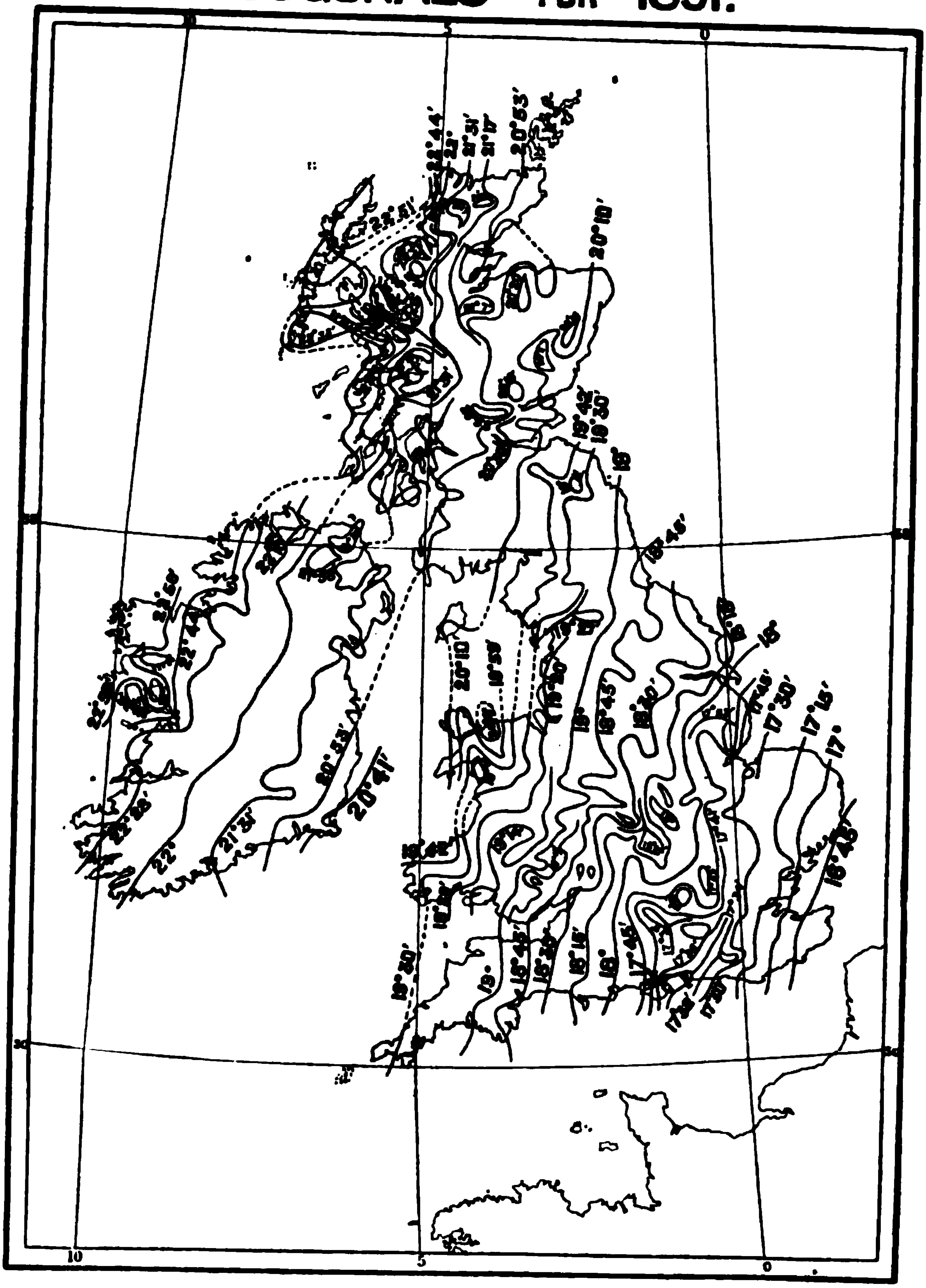


FIG. 426.

which the actual measured declination is the same. Here the effects of local disturbances have not been eliminated by lumping the stations together in groups so as to eliminate the effects of these disturbances, at any rate when they do not affect any large area.

From the amount of the differences between the values of the elements as obtained from the smooth curves and the actual measured values, an idea of the position and extent of the magnetic masses which cause these differences has been made by Professor Rücker, so that the value of the magnetic field on the surface of the earth is employed to get an idea of the geology of the portions of the earth's crust below the actual surface layers.

434. Continuous Magnetic Records.—In a certain number of observatories a continuous record of the values of the magnetic elements is kept by means of self-recording instruments. The records are obtained by means of the trace left by a spot of light reflected from a mirror attached to a magnet on a sheet of photographic paper, which is kept in uniform movement by means of clockwork. In the case of the declination, the mirror is simply attached to a magnet which is suspended by a long fine thread, so that it can turn freely about a vertical axis, and so, by always setting itself in the magnetic meridian, gives a record of the changes that take place in the position of this meridian, that is, shows the changes in the declination.

The changes in the horizontal force are recorded by means of a magnet which is suspended by a bifilar suspension (§ 119). The top of the bifilar is turned till the magnet sets itself at right angles to the magnetic meridian, under which circumstances the earth's field exerts a turning couple on the magnet equal to MH (§ 425), this couple being balanced by the couple due to the bifilar. If the value of the horizontal force H alters, the couple due to the magnetic forces alters also in the same proportion, and the magnet turns about a vertical axis till the couple due to the bifilar becomes equal to the new couple due to the magnetic forces. Changes in the declination will, however, not affect the position of the magnet, since it is at right angles to the magnetic meridian.

Since no satisfactory way of recording the changes that take place in the dip has been devised, it is usual to record the changes in the vertical force. For this purpose a magnet is balanced on knife edges in such a way that it is in an approximately horizontal position. If, say, the vertical force decreases, then the downward force acting on the north pole and the upward force acting on the south pole both decrease, and hence the north pole of the balanced magnet rises and the south pole falls, just as when, in a balance, the load of one pan is increased and that of the other is decreased. The motions of such a balanced magnet will therefore indicate the changes that take place in the vertical force, and since the magnet with the bifilar suspension gives the changes that

take place in the horizontal force, the changes in dip and in the total force can be immediately calculated from the records given by the two instruments.

435. Diurnal Range.—As a result of a study of the records of the magnetographs, as the self-recording magnetic instruments are called, it is at once evident that the values of the magnetic elements undergo small daily changes in value, the magnitude of this diurnal range depending on the position of the place and the time of year. The form

FIG. 427.

of the diurnal range curves for Kew for the summer months are shown in Fig. 427, which gives the variation of each element from its mean value for the whole twenty-four hours. The fact that the curve is above the zero line means that the corresponding element is greater than its mean value.

436. Annual and Secular Change.—In addition to the diurnal range, the magnetic elements undergo a periodic change of which the period is a year, which is called the annual range.

Slow changes of which, if they are periodic, the period must be many centuries, also take place in the values of the elements, and these are called secular changes. In Fig. 428 the change in the value of the declination at London during the last three hundred years is shown by means of a curve. It will be seen that the declination attained a maximum westerly value in 1810, while in 1660 the declination was zero, so that in that year the agonic line passed through London.

A very elegant method of showing the changes due to the secular variation has been introduced by L. A. Bauer. If we suppose a magnet

suspended in such a way that it is free to set itself parallel to the lines of force of the earth's field, then, owing to secular change in the declination and in the dip, the north pole of the magnet would describe a curve in space. The form of the curve in the case of a magnet in London is shown

Fig. 428.

in Fig. 429. From this curve, and similar ones drawn for other places, Bauer was able to show that the north end of such a freely suspended needle describes a curve such that, to an observer situated at the centre of the needle, the curve is described in the same direction as that in which the hands of a watch move. The form of the curve given in Fig. 429

Fig. 429.

seems also to indicate that the curve described by the pole of the needle will be closed, the time taken for the needle to complete a whole cycle being about 470 years.

437. Magnetic Storms.—In addition to the regular changes in the magnetic elements which we have been considering, sudden disturbances of these elements sometimes occur, which are often, especially when the

phenomenon called the aurora borealis is seen, of considerable magnitude. The character of such magnetic storms, as they are called, is shown by the copy of the photographic trace of the self-recording declination magnetograph of Greenwich Observatory during a magnetic storm and

Fig. 430.

also during an ordinary quiet day, reproduced in Fig. 430. The cause of these magnetic storms has not yet been discovered, although there seems to be some connection between them and the condition of the sun, for whenever there are a large number of spots on the sun there always seem to be a number of magnetic disturbances.

PART II—ELECTRO-STATICS

CHAPTER III

ELECTRO-STATIC ATTRACTION AND REPULSION—COULOMB'S LAW

488. Fundamental Experiment.—Thales, who lived about the commencement of the Christian era, discovered that amber when rubbed acquires the property of attracting light bodies, such as pieces of pith or cork. Towards the end of the sixteenth century Gilbert showed that this property was also possessed by other bodies, such as wax, sulphur, and glass. All such phenomena are studied in the science of electricity, the name being derived from the Greek name for amber.

A body which has acquired this property of attracting other bodies, the attraction considered being of course different from the gravitational attraction which all bodies exert one on the other, is said to be electrified, or to possess electrification. Electrification, unlike mass, is not a fundamental property of matter, since under ordinary circumstances matter is unelectrified, and it is only after the electrification has been produced by certain causes, which we shall examine in detail later on, that it becomes electrified.

The most usual manner of causing the electrification of a body is that referred to above, namely, friction with a suitable rubber. Thus a stick of sealing-wax, when rubbed with a dry piece of flannel, becomes electrified, as also does a rod of glass when rubbed with silk.

489. Conductors and Non-Conductors.—All substances may be roughly divided into two classes, called conductors and non-conductors. In a conductor the electrification spreads all over the body, so that if one point of the body is by any means electrified, this electrification immediately spreads all over the body. In the case of a non-conductor, or insulator, as such bodies are also called, the electrification does not spread in this way, but remains in the neighbourhood of the point where the electrification took place.

The best conductors are the metals and solutions of most salts in water, while the best non-conductors are ebonite, glass, shellac, sulphur, paraffin, sealing-wax, and silk. There is, however, no hard and fast line of demarcation between the two classes, for such bodies as dry wood

and paper have intermediate properties, and are sometimes called semi-conductors. In the study of electricity it is of much importance to have a good non-conductor, for by this means we are able to support a body in such a way that any electrification communicated to it will not spread to neighbouring bodies through the support. Although no body is known which is a perfect insulator, yet glass, particularly when it has been boiled in water and is then kept in a dry atmosphere, paraffin, and fused quartz are sufficiently good insulators for all practical purposes. When a body is supported on an insulating stand, we shall speak of it as being insulated.

440. Two Kinds of Electrification.—If a rod of sealing-wax is electrified by rubbing with flannel, and is then suspended by an insulating thread, such as silk, and a second rod of sealing-wax is also electrified in the same way and brought near the first, they will repel each other. We have here a case then of two electrified bodies repelling one another. In the same way, if two rods of glass are electrified by being rubbed with silk, and one of them is suspended by the silk thread and the other brought near, repulsion will take place. If, however, a rod of sealing-wax, electrified by friction with flannel, is brought near the glass rod, which has been electrified by friction with silk, the two will attract one another. We thus see that we have here to do with two kinds of electrification, in the same way that in the case of magnets we had to do with two kinds of poles. The kind of electrification that is developed in glass when it is rubbed with silk is distinguished by being called positive electrification, while the kind of electrification produced in sealing-wax by friction with flannel is called negative.

We may then state the law of electrical attraction and repulsion as follows: Bodies electrified in the same manner repel one another, while bodies electrified, one positively, and the other negatively, attract one another.

Whenever electrification of one kind is produced in any way, electrification of the opposite kind is also produced at the same time. Thus in the case of the glass electrified by friction with silk, while the glass will attract a negatively electrified rod of sealing-wax, the silk used to rub the glass will repel the sealing-wax, thus indicating that the silk has become negatively electrified.

The kind of electrification developed in a body depends on the nature of the body with which it is rubbed; thus while glass becomes positively electrified when it is rubbed with silk, it becomes negatively electrified when it is rubbed with a cat's skin. The kind of electrification produced is also dependent on the state of polish of the surface, on the temperature, &c.

441. The Gold-Leaf Electroscope.—In order to study the sign, and to a certain extent the magnitude of the electrification produced in a given body, the instrument shown in Fig. 431, and called the gold-leaf

electroscope, is often convenient. It consists of a glass flask, in which hang two gold leaves, CC, which are in conducting communication with a metal disc, A, by means of a metal rod, the rod being insulated by a coating of shellac, D. When the metal disc or cap of the instrument is put in conducting communication with an electrified body, the gold leaves both become electrified with the same kind of electrification as the body; and since two bodies electrified in the same way repel one another, they diverge as shown in the figure, the amount of the divergence being a rough measure of the amount of the electrification of the body.

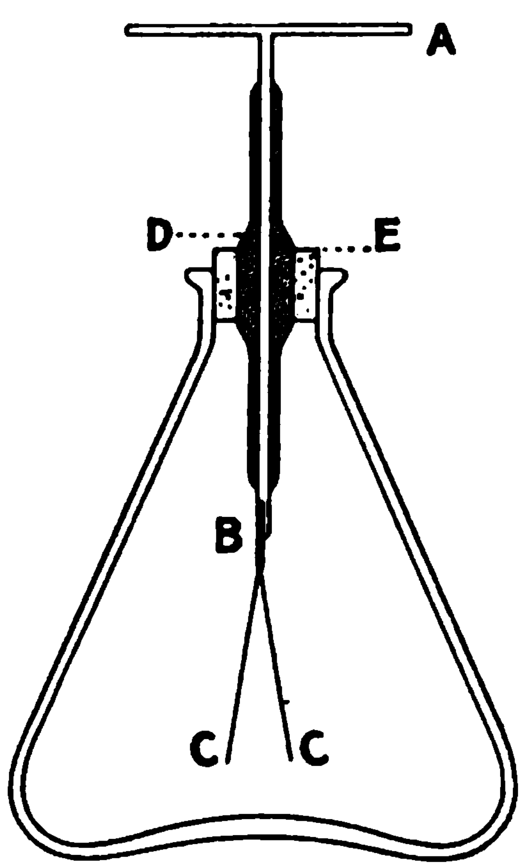


FIG. 431.
(*From Watson's "Elementary Practical Physics."*)

442. Electrification by Induction.—If an electrified body is brought near the cap of a gold-leaf electroscope, it will be found that the leaves diverge, showing that they have become electrified before the electrified body has come into conducting communication with the cap. On the removal of the electrified body the leaves again collapse, showing that they have lost the electrification they possessed when the electrified body was near. This electrification, caused by the *proximity* of a charged body, is said to be produced by induction.

If the inducing body is charged positively, the part of the insulated body nearest to the inducing charge will be negatively electrified, while the part furthest from the inducing charge will be positively electrified. That this is so can easily be shown by means of a small piece of metal attached to an insulating handle, and called a proof-plane, which is brought into contact with different parts of the body on which the induced charges are produced. The sign of the charge carried away by the proof-plane, after contact with any given part of the body, can be found by means of the gold-leaf electroscope. In this way it can be shown that whenever an insulated conductor is placed in the neighbourhood of a charged body, the conductor will become electrified by induction, the electrification at the end nearest the charged body being of the opposite kind to that of the charged body, while the electrification on the end furthest from the charged body is of the same kind as that of the inducing charge. If, while an insulated conductor is in the neighbourhood of a charged body, so that it is charged by induction, it is placed in conducting communication with the earth, the electrification of the same kind as that of the inducing charge will be destroyed. If the connection with earth is now broken, and the inducing charge is then removed, it will be found that the conductor is now electrified with

the opposite kind of electrification to that of the inducing body. In the case when the inducing charge is removed before the conductor has been put to earth, the reason why, on the removal of the inducing charge, the conductor was unelectrified was that the two kinds of electrification, produced in equal quantities by the induction, neutralise each other. We shall return to the consideration of the subject of induction after we have dealt with the quantitative measurement of electrification.

443. Coulomb's Law.—By means of the torsion balance, Coulomb was able to show that the force exerted on one another by two charged conductors is directly proportional to the product of their charges, and inversely proportional to the square of the distance between the bodies.

Hence, as in the case of the unit magnetic pole, we may define the unit electrification or charge as such that if two small bodies, each charged with a unit, are placed at one centimetre apart in air, the force they will exert on one another will be one dyne. The reason the medium air is specified is that, as we shall see later, the force exerted between two charged bodies depends on the nature of the medium which fills the space between them, while the reason the bodies on which the charges are supposed to exist are taken as small is that if the bodies were of appreciable magnitude the distribution of the electrification would be altered by the action of the one charge on the other.

Suppose then we had two points, charged with e and e' units of electricity respectively, placed at a distance r apart in air, the force, F, which they would exert one on the other, due to their electrification, will be given by the equation—

$$F = \frac{ee'}{r^2}.$$

The force will be an attraction if the charges e and e' are of opposite sign, and a repulsion if they are of the same sign. In the case of the unit of electrification, as we shall see later, we meet with a case where there are two separate relations commonly employed to connect the quantity to be measured with the fundamental units (§ 8). Hence, in order to distinguish the unit as defined above, which depends for its definition on the force exerted between two charged bodies, and another unit which we shall consider later, and which depends for its definition on another physical property of a charged body, the unit above defined is called the *electro-static unit* of quantity of electricity or charge.

CHAPTER IV

THE ELECTRICAL FIELD

444. Electrical Lines of Force.—If a small body, charged with the unit positive charge, is brought into the neighbourhood of a charged body, this unit charge will be acted upon by an electrical force, which at every point of the space surrounding the charged body will have a definite magnitude and direction. As in the case of magnetism, a line, such that its direction at every point is the same as the direction of the force acting on the unit charge when placed at the point, is called a line of force. The direction in which a line of force is supposed to run is the direction in which a small positively electrified body would tend to move. Hence a line of force will always start from a body which is positively electrified and end on a body which is negatively electrified.

If we take an area on the surface of a positively electrified body, such that this portion of the surface contains a unit of electricity, and, starting from all points on the curve bounding this area, draw the lines of force, these lines of force will enclose a tube-shaped space which is called a tube of force. Since each of the lines of force must terminate on a negatively electrified body, we see that every tube of force must end on a negatively electrified body, and it can be shown that the quantity of electricity on that portion of the surface enclosed by the tube of force will be a unit of negative electricity. By means, therefore, of tubes of force, we can indicate the distribution of the electrification on the surface of a charged body, for the greater the charge the smaller will be the cross-section of the tubes of force, and hence the larger the number of them which will leave each square centimetre of the surface of the charged body. Since it would be rather inconvenient to draw a series of tubes, it is usual to suppose that a single line of force is drawn along the axis of each unit tube of force, that is, that from the centre of each element of the surface of the positively electrified body on which the unit quantity of positive electricity exists we draw a line of force. Under these circumstances the number of these lines of force which leave the surface of the charged body will represent the charge on the surface. If a body is charged with e units of electricity, then e lines of force will leave the surface of the body, while if the body is charged with e units of negative electricity, e lines of force will terminate

on the surface of the body. It will thus be seen that if a body is charged with e units of positive electricity, so that e lines of force leave the body and must terminate on a negatively charged body, somewhere or other there must necessarily exist e units of negative electricity. This negative charge may, however, be so far removed from the spot where we are making our experiments that it does not in any way affect the results, and hence we are able to perform experiments in which we practically have only to deal with one kind of electrification. Here we have a marked difference between magnetism and electricity, for in the case of magnetism we are unable to obtain a body which has only one pole, and so cannot deal with a single pole.

The space in the neighbourhood of electrified bodies in which electrical phenomena, such as attraction, are exhibited is called an electrical field. A field in which the force acting on a small electrified body is everywhere the same both in magnitude and direction is called a uniform field. In a uniform field the lines of force must be everywhere parallel, and therefore the tubes of force must everywhere have the same cross-section.

The quantity of electrification on the unit of area of the surface of an electrified body is called the surface density of the electrification, and, as we have seen, the number of lines of force which leave or terminate on the unit of area of the surface is also equal to the charge on the unit area. Hence the surface density may also be defined as the number of lines or tubes of force which leave or terminate on the unit area of the surface of the electrified body. In the case where the electrification of the body is not uniform, the surface density at a given point is defined, as in the case of other variable quantities, as the quantity of electricity on a small element of surface surrounding the given point divided by this area.

The lines of force in the case of two small bodies, one of them positively and the other negatively electrified, and placed at a very great distance from all other conductors, so that all the lines of force which leave the positively electrified body terminate on the negatively electrified body, are shown in Fig. 432,[1] while in Fig. 433 the lines of force in the case where the two bodies are electrified with the same kind of electrification are shown.

As in the corresponding case in magnetism, we may account for the attraction which takes place in the one case, and the repulsion in the other, if we suppose that there exists a tension along the lines of force, and that something of the nature of an hydrostatic pressure acts at right angles to the direction of the lines, so that they repel one another.

When the electrical charge on any system of conductors alters its distribution, we may consider that each unit of the charge, as it moves over the surface of the conductors, drags the end of its tube of force after

[1] The lines of force are symmetrical about the line joining the charges, and so to save space only half are shown.

it, but that, on account of the tension acting along the tube, the tendency is for the tube to become as short as possible. When the two conductors, on which any given tube terminates, are separated by a non-conductor, the tube of force cannot shorten indefinitely, for the ends of the tube cannot leave the conductors. If, however, the two conductors are placed in conducting communication, say by being joined by a wire, the ends of the tube can now move along this wire, so that the tube can shorten indefinitely, and ultimately vanish.

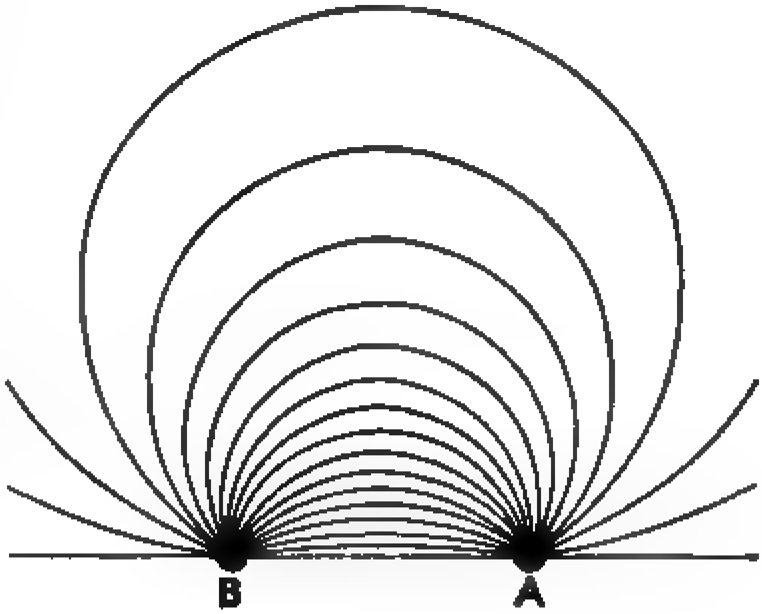


Fig. 432.

Thus by supposing that not only does the tension along the lines of force give rise to a mechanical force acting on the matter on which the electrification exists, but also that this tension causes the electricity of the two opposite kinds which exist at the two ends of the line of force to tend to approach each other, and can only be kept apart by the interposition of a non-conductor, we shall be able to explain how it is that one of the kinds of electricity produced by induction remains on the body when the latter is put to earth, while the other kind of electrification escapes to earth.

Fig. 433.

In Fig. 434 let A represent the inducing body, which we may suppose charged with positive electricity, and B be an insulated conductor which is electrified by induction by A. Then some of the lines of force (shown by the full lines) which leave A will terminate on B, and B will therefore be negatively electrified at the part where these lines meet the surface. In addition a number of lines of force will leave B, and terminate on surrounding conductors, such as the walls of the room in which the two

bodies are placed. The part of B where these lines leave the surface will be positively electrified, the corresponding negative charge being on the walls. The lines of force which stretch from A to B, by their tension, cause the negative charge on B to accumulate on the side next A. The whole charge does not accumulate at the nearest point, however, because of the mutual repulsion which the lines of force exert on one another. It

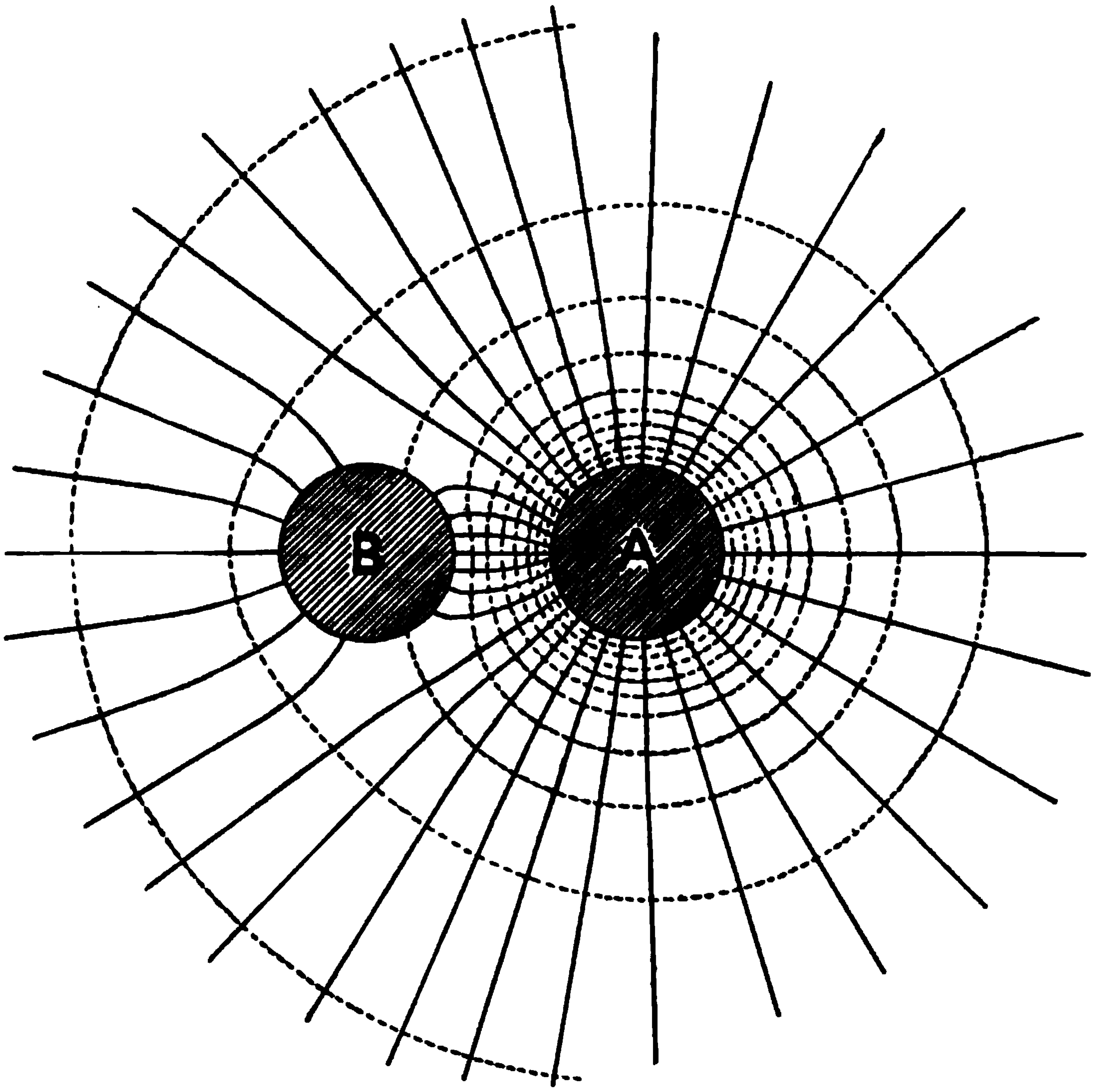


FIG. 434.

is owing to this repulsion between the lines that the lines leaving the body B accumulate at the other end.

When the body B is put in conducting communication with the earth, *i.e.* with the bodies on which the lines of force which leave it terminate, owing to the action of the tension on the electrification itself, the latter will escape, but the negative electrification corresponding to the lines of force which leave A and terminate on B will not be able to reach A, since these two bodies are not in conducting communication. The distribution

of the charges will then be as shown by the lines of force in Fig. 435, where there are no lines of force leaving B, indicating that the charge on B is everywhere negative.

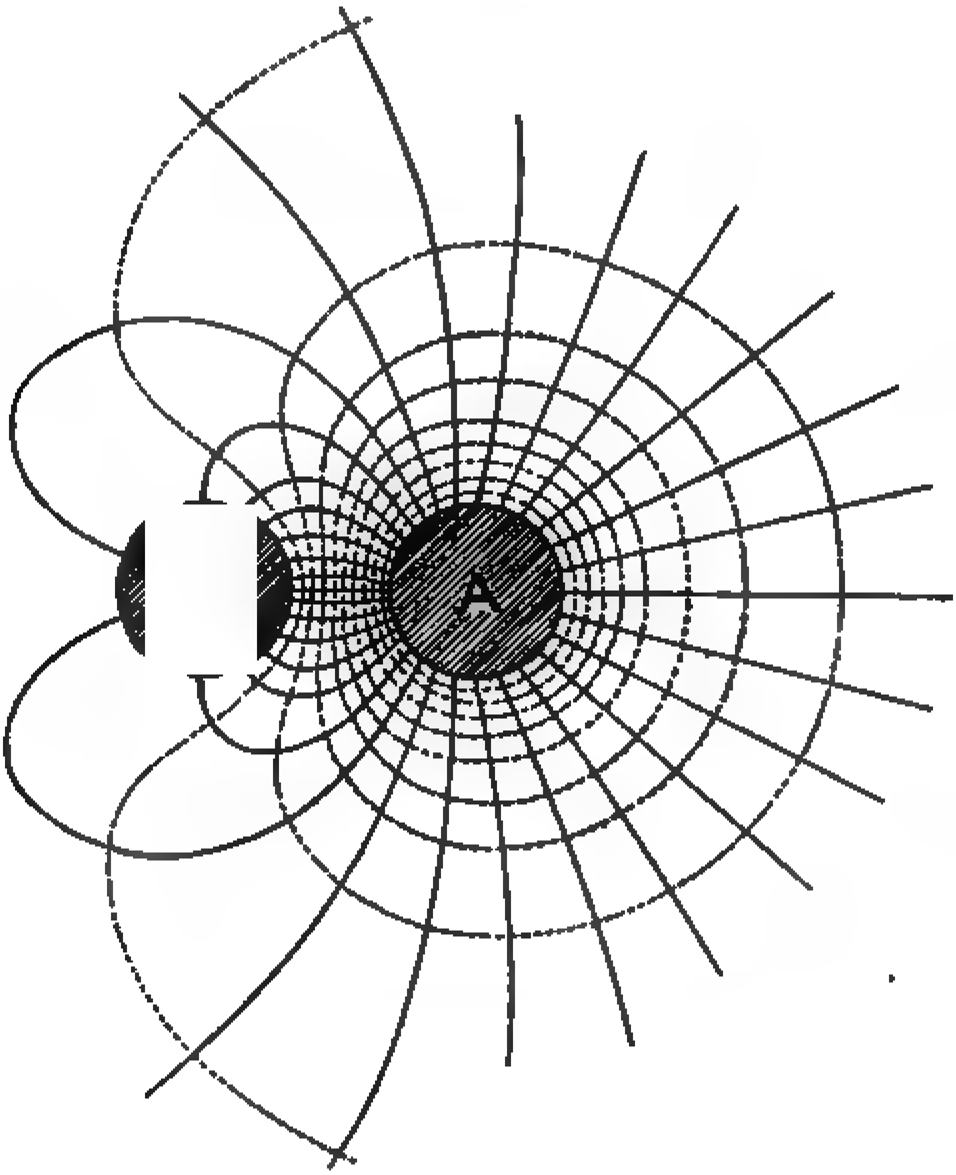


FIG. 435.

445. Faraday's Ice-Pail Experiment.—The production of electrification by induction can be very clearly investigated by means of the arrangement shown in Fig. 436. Faraday, to whom these experiments are due, used, in place of the hollow metal sphere A, a metal ice-pail, and owing to this circumstance it is generally known as Faraday's ice-pail experiment. The hollow metal sphere has an opening B, through which a small charged sphere D can be lowered into the interior. The hollow sphere is supported on an insulating stand, and is connected with an electroscope C. Let A be unelectrified, and suppose that we introduce the small sphere D, which is suspended by an insulating thread, and is

charged with Q units of positive electrification. Before D was introduced within A, the tubes of force which start from the charged sphere terminated on the walls of the room and on neighbouring conductors. When, however, D is introduced within A we shall have two sets of tubes, one set starting from D and terminating on the inside of A, and the other starting from the outside of A and terminating on the walls of the room. In other words, the conducting shell A has divided each of the tubes of force which leaves the charged body D into two parts.

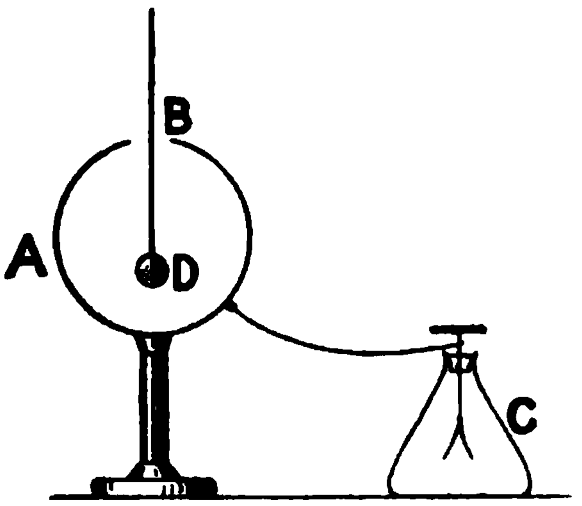


FIG. 436.

The tubes of force which leave the outside of A correspond to a positive electrification, and the magnitude of this electrification is indicated by the magnitude of the divergence of the gold leaves of the electroscope. If now the vessel A is put in conducting communication with the earth, the tubes of force which stretch from its outer surface to the walls will be able to contract, the ends running along the conductor used to put A to earth. The tubes which terminate on the inside surface will, however, not be able to shrink, for the sphere D is not in conducting communication with the vessel A. Next let the vessel A be again insulated, and then lower the sphere D till it touches the inside of A. Now the tubes of force between D and the inside of A are able to shrink and vanish. When the vessel A was put to earth the leaves of the electroscope collapsed, showing that the vessel A had lost its free charge, and they remain collapsed even when the charged sphere is allowed to touch the inside of the vessel A. This therefore shows that the positive charge on D is exactly equal to the induced negative charge on A. Hence the charge induced on the inside of A is $-Q$. Also, since there were Q tubes of force leaving D, and each of these tubes must have terminated on the inside of A, that is, on a surface on which the charge is $-Q$, we see that this experiment proves that the charge on the portion of the surface on which each tube terminated was a unit of negative electricity.

Next remove D and again charge it with Q units of positive electricity, that is, give it the same charge as before, and again introduce it within the vessel A. If now the sphere D be lowered till it touches the inside of A, it will be found that the separation of the leaves of the electroscope remains unaltered. We have now communicated a charge $+Q$ to A, and since this produces the same deviation of the electroscope leaves as did the induced charge produced by the body charged with $+Q$ units placed inside, we see that the positive charge produced by induction is equal to Q units. But the previous experiment showed that the negative charge

produced by induction on the inside of A is also equal to Q units. Hence we see that the positive and negative charges produced by induction are equal. Thus our assumption that every tube of force starts from a portion of a conductor on which there is a unit positive charge, and terminates on a portion on which there is a unit negative charge, is justified.

Since every tube of force must have both a beginning and an end, it therefore follows that to every positive charge there must exist an equal negative charge, this charge being situated on the bodies on which the tubes of force which leave the positively charged body terminate. From this it follows that whenever we, by any means whatever, give a charge of one sign to a body an equal and opposite charge must at the same time be produced. This deduction may be proved experimentally by means of Faraday's ice-pail experiment, for if a small piece of sealing-wax is attached to an insulating handle and introduced within the vessel A, and is electrified by rubbing with a flannel pad attached to a second insulating handle, the electroscope will be unaffected. The reason is that the charges produced on the sealing-wax and the flannel are equal and opposite, and therefore they induce equal and opposite charges on the outside of the vessel A. On removing either the sealing-wax or the flannel the electroscope will be affected, for now the inducing charge inside A is all of one sign, and hence so also is the charge induced on the outside.

We can by this arrangement give the vessel A a charge, the magnitude of which is any given number of times the magnitude of some given charge. Thus suppose we have a negatively charged sphere, and that we bring the sphere D to within, say, six inches and then put it momentarily to earth. In this way D will obtain a charge of positive electricity of magnitude Q, say. Next introduce D inside A, and let it touch the bottom. In this way we shall communicate a charge of Q to A. If now D is again brought to within six inches of the negatively charged sphere earthed and then introduced within A as before, a further charge of Q will be communicated to A, so that the total charge is 2 Q, and so on.

446. Difference of Potential.—If two conductors, one of which is charged positively and the other is charged negatively, are put in conducting communication, their state of electrification will become changed, so that if they originally possessed equal charges they will both, after being connected, exhibit no signs of electrification. If the charge on one was greater than that on the other, then, after being connected, the sign of the charge on the two will be the same as the sign of the charge which was originally the greater, while the sum of the charges now possessed will be equal to the difference of the two original charges. If, however, two bodies, each of which is charged with electricity of the same kind, are put in conducting communication, it does not follow that the charge on the body which was originally electrified with the larger charge will be decreased and that of the other

increased, for a small sphere charged with one unit of positive electricity, when put in conducting communication with a sphere, of which the radius is three times that of the other and which is charged with two units of positive electricity, will lose electrification. There must evidently, therefore, be some other condition besides the magnitude of the charge which decides whether, when two charged bodies are put in communication, the charge of one or other of them becomes increased.

Two conductors are said to be at different potentials if, when they are put in conducting communication, the distribution of electrification on the conductors changes. The body on which the positive electricity *decreases* is said to be at the higher potential.

This idea of electrical potential is of the same nature as the idea of temperature in the case of heat, or of level in the case of the flow of water in a pipe, for, as we have seen, heat always flows from a body at a higher temperature to a body at a lower temperature, and water only flows from places at a higher level to places at a lower level.

The difference in potential between two charged conductors is measured by the work that would have to be done on a small body charged with a unit of positive electricity when the body is moved from the immediate neighbourhood of the conductor at the higher potential to the immediate neighbourhood of the conductor at the lower potential.

For all practical purposes the measure of the available energy of a waterfall is known if the available head and the quantity of water which passes in a second are known, for the variation in the value of the acceleration due to gravity (g) is comparatively small. It would be quite otherwise, however, if the value of g varied to any great extent from one place to the other on the surface of the earth. Thus suppose that we had to do with two waterfalls in which the quantity of water which passed per second was the same, but the fall was different and the value of g was twice as great at one place as at the other. Then the work which could be obtained from the unit mass of water as it passed from the top to the bottom of the fall would be gh_1 in the one case and $2gh_2$ in the other. Hence, as far as the quantity of energy available is concerned, the height through which the water falls, that is, the difference in level between the water above and below the fall, is not a measure of the value of the fall. If, however, we measured this difference of "level" by the quantity of work which must be done to raise unit mass of the water from the bottom of the fall to the top, then the available energy of any fall would be simply obtained by multiplying this quantity by the quantity of water which passes over the fall in a unit of time. Now although, as has been mentioned above, the changes in g are so small as to make it quite unnecessary to adopt, in the case of waterfalls, any such device, yet it will be seen why the method adopted for measuring the difference of potential between two charged bodies is quite a reasonable one.

When considering the absolute scale of temperature in § 261 we used a very similar method, for the difference in temperature between two bodies was measured by the work which could be done by a reversible engine when working between these two temperatures, and taking a given quantity of heat from the hotter body. Thus in this case also a quantity of work is used as a measure of the difference of the quantity (temperature) which decides in which direction heat will flow when two bodies are placed in thermal communication, and is, therefore, analogous to potential in the electrical problem.

The amount of work done on the unit of positive electricity as it is carried from the neighbourhood of one charged body to that of the other is the same, whatever the path by which it is moved. If it were not, so that it were possible to pass from a point A (Fig. 437) to another point B at a lower potential, in such a way that the work w_1 done on the unit charge when taken along the path ACB was greater than the work w_2 done when the unit is moved along the path ADB, then by taking the body with the unit charge from A to B by the path ACB, and bringing it back by the path BDA, the whole system would have performed a cycle, for the initial and final states are the same, while an amount of work equal to $w_1 - w_2$ would have been done without the supply of any external energy. This being contrary to the doctrine of the conservation of energy, it follows that w_1 must be equal to w_2, that is, the work done when the unit charge is carried from A to B must be independent of the path by which it is carried.

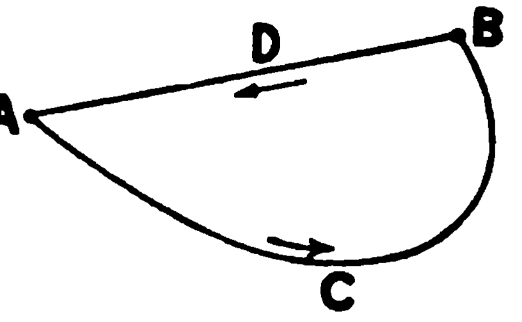


FIG. 437.

We have defined the difference between the potential of two points and shown how it is measured, and we have now to choose some fixed potential as the zero of potential. The potential of the earth is usually taken as the zero of potential, so that the potential of a positively electrified body is positive, and that of a negatively electrified body is negative, for a positively electrified body will repel a body charged with a unit of positive electricity, and so work will be done on the unit charge as it is moved from the electrified body to the earth, while if the body is negatively electrified, work must be supplied to move the unit charge from the electrified body to the earth.

447. Equipotential Surfaces.—An equipotential surface is a surface such that the potential of all points upon it is the same. No work is therefore done when a charged body is moved along a path which lies in an equipotential surface. It follows at once that the lines of force must always cut an equipotential surface at right angles. If a line of force did not cut an equipotential surface at right angles, then the force which acts in the direction of the line of force can be resolved into two

components, one along the surface and the other normal to the surface. If an electrified body were placed at the point where the line of force cuts the equipotential surface, it would be acted upon by the component parallel to the surface, and if it were moved in the direction in which this component acts, work would either be done on or by the electrified particle. But by the definition of an equipotential surface no work is done when a charged body is moved from one point of such a surface to any other point on the surface. Hence it follows that the component of the force parallel to the surface of the equipotential surface must be zero, or, in other words, that the direction of the line of force must be perpendicular to the surface at the point where it cuts the surface.

Fig. 438.

Since in the case of a conductor the electrification is not prevented from spreading itself over the surface of the body, no change in the distribution of the electrification would take place by connecting any two points of the surface by a conducting wire, and so all parts of the surface must be at the same potential. The surface of a conductor must therefore be an equipotential surface, and hence the lines of force must always cut the surface of a conductor at right angles to the surface.

In Fig. 438 the lines of force and the equipotential surfaces for a positively charged body A are shown, the trace of the equipotential surfaces being shown by the dotted lines. If an insulated uncharged

conductor B is placed in the neighbourhood of the charged conductor, this conductor will become electrified by induction. Now if the conductor B could be brought near the charged body A without producing any change in the distribution of the charge on the conductor or changing the state of the electrical field in the space now occupied by the conductor, that is, if the lines of force and the equipotential surfaces were to remain as in Fig. 438 *after* the introduction of the conductor, then those parts of the conductor B furthest from A would be at a lower potential than the parts nearer A. Hence, since it is impossible for different parts of a conductor to be at different potentials so long as the electrification is not changing, some change in the electrical conditions must take place so as to raise the potential of the more distant parts of the conductor B, or lower the potential of the nearer parts. This will occur if the more distant parts become positively electrified and the nearer parts negatively electrified, for under these circumstances a greater repulsive action will be exerted on a unit of positive electricity when placed near to the further surface of B, and hence a greater amount of work will be done on this unit while it is being moved from this position to the neighbourhood of the earth. In the same way, less work will have to be done to move a unit charge from the near side to the neighbourhood of the earth, so that the potential of the near side will be reduced by the presence of the induced negative electrification. This lowering of the potential on the near side of B, itself involves a lowering of the potential of the near side of the conductor A, and hence also of the far side. This lowering of the potential of the far side is produced by the accumulation of the positive electrification of A on the side near B.

The form of the lines of force and of the equipotential surfaces under the new conditions is shown in Fig. 434. It will be seen that the change in the distribution of the charge on A, as well as the distribution of the induced charge on B, is such that the surfaces of the two conductors are equipotential surfaces. If the insulated conductor is earthed, then the electrification on both conductors is altered, but in such a way that, as shown in Fig. 435, the surfaces of the conductors remain equipotential surfaces. We thus see that the fact that the distribution of the electrification on the body, when placed in the neighbourhood of a charged body, is not uniform is not inconsistent with the surface of the conductor being an equipotential surface, but is in fact the distribution which, in conjunction with the inducing charge, insures the fulfilment of this condition.

If A and B are two points on a line of force, they must necessarily be at different potentials. Let the potential of A be V_1, and that of B be V_2 (V_1 being greater than V_2); then if a small body carrying the unit charge of positive electricity is moved along the line of force from A to B, the work done will be equal to $V_1 - V_2$, for the difference in potential between two points is measured by the work done on the unit charge when it is

moved from one point to the other. If the points A and B are very close together, the force F, which acts on the unit charge in the direction from A to B, as it is moved from A to B, may be supposed to remain constant, and to be equal to the average force that acts. The work done by the unit charge as it is moved from A to B will therefore be equal to Fs, where s is the distance from A to B, measured along the line of force, that is, the distance through which the charge is moved along the line of action of the force F. Hence the work done *on* the unit charge is $-Fs$. Equating the two expressions we have now obtained for the work done on the unit charge when it is moved from A to B, we get

$$V_1 - V_2 = -Fs,$$

or

$$F = -\frac{V_1 - V_2}{s}.$$

Now the expression on the right-hand side of this equation is the difference of potential between the two points divided by the distance between the points measured along a line of force, or, in other words, is the rate of change in the potential along the line of force at the points A and B, which are by supposition very close together. Hence the force which acts on a unit charge of positive electricity, when placed in an electrical field, is equal to minus the rate of change of the potential along the line of force at the given point. If the force acting on the unit charge is constant, it follows that the rate of change of the potential must also be constant. Hence in a uniform electrical field (§ 444) the rate of change of the potential in the direction of the lines of force must be constant, and thus the length of a line of force intercepted between two consecutive equipotential planes will be the same, if the difference of potential between consecutive equipotential planes is itself constant.

448. Electrification Confined to the Surface of a Conductor.—If a hollow conducting vessel is electrified, the whole of the charge is confined to the outside surface of the conductor. That this is so may be shown by touching the inside surface with a proof-plane, then removing the proof-plane, taking care not to touch the sides of the orifice of the charged vessel, when it will be found on testing the proof-plane that it has not carried away any charge. Another way of proving that the charge is entirely on the outside is to lower a small charged sphere, attached to an insulating thread, into a hollow conductor, and allow it to touch the inside of the conductor, and then withdraw it, when it will be found to have completely lost its charge. When the charged sphere was allowed to touch the inside of the hollow conductor, it, for the time being, formed part of this conductor; and since it entirely lost its charge, we can infer that the charge on a conductor is entirely confined to the outside surface.

449. Force Exerted on a Charged Body placed within a Hollow Charged Conductor.—Since the surface of a charged con-

ductor is an equipotential surface, the whole of the space within must be at the same potential, so long as there are no charged bodies within the conductor. For, suppose that within the surface of the conductor there were an equipotential surface corresponding to a higher potential than the potential of the surface of the conductor, then there would be lines of force running everywhere to the outer equipotential surface from this inner one; and since these lines of force must of necessity start from a positively electrified body, it would follow that there must be a positively electrified body within the conductor, which is contrary to our original supposition. In the same way it would follow that, if there existed an equipotential surface of lower potential than that of the surface of the conductor, there must be a negatively electrified body within the conductor. We are therefore led to the conclusion that there can be no point within a closed conductor at a different potential from that of the surface, unless there are charged bodies within the conductor.

Since the strength of an electrical field is equal to minus the rate of change of the potential, it follows that if the potential is constant, that is, if its rate of change is zero, there will be no electrical force exerted within the conductor. We thus see that it follows, from the fact that the charge of a conductor is confined to the outside surface, that there is no force exerted within a charged conductor; and it can be shown that this condition can only be fulfilled if Coulomb's law (§ 443), that the force exerted between two charged bodies varies inversely as the square of the distance, is true.

Take the case of a uniformly electrified sphere, and suppose we require to find the force at a point P (Fig. 439). Through P draw a series of lines forming a cone with P as vertex, and intersecting the surface of the sphere in the small areas s and S. Then it can be shown[1] that, if r and R are the distances of these areas from the point P, then s is to S as r^2 is to R^2. Hence, as the density σ of the charge on the sphere is uniform, the charges on the areas s and S are proportional to these areas, that is, in the ratio of r^2 to R^2. If now the force exerted is inversely as the square of the distance, the ratio of the forces exerted at P

[1] Since the tangents at s and S make equal angles with the chord sPS, they must make equal angles, θ, with lines drawn through the points a and b at right angles to the chord. If ω is the solid angle of the cone at P, the area of a cross section of this cone made by a plane, ac, at right angles to the axis of the cone is ωr^2. Hen e the area intercepted by the cone on a plane inclined at an angle θ to the axis is $\omega r^2/\cos\theta$. If the angle ω is small, the area intercepted on the tangent plane is the same as the area intercepted on the sphere, and so the area, s, of the surface of the sphere intercepted by the cone is $\omega r^2/\cos\theta$. In the same way the area, S, intercepted by the cone is equal to $\omega R^2/\cos\theta$. Hence

$$\frac{s}{S}=\frac{\omega r^2}{\cos\theta}\Big/\frac{\omega R^2}{\cos\theta}=\frac{r^2}{R^2}.$$

by the charges on the surfaces s and S will be as $\frac{R^2}{R^2} : \frac{r^2}{r^2}$. Hence we see that, if the inverse square law holds, the forces exerted at the point P by the charges on the portions of the surface of the sphere intercepted by the cone are equal and opposite, so that they neutralise each other. The same will hold for the charges on the portions of the surface intercepted by any other cone drawn through P ; and since the whole surface of the

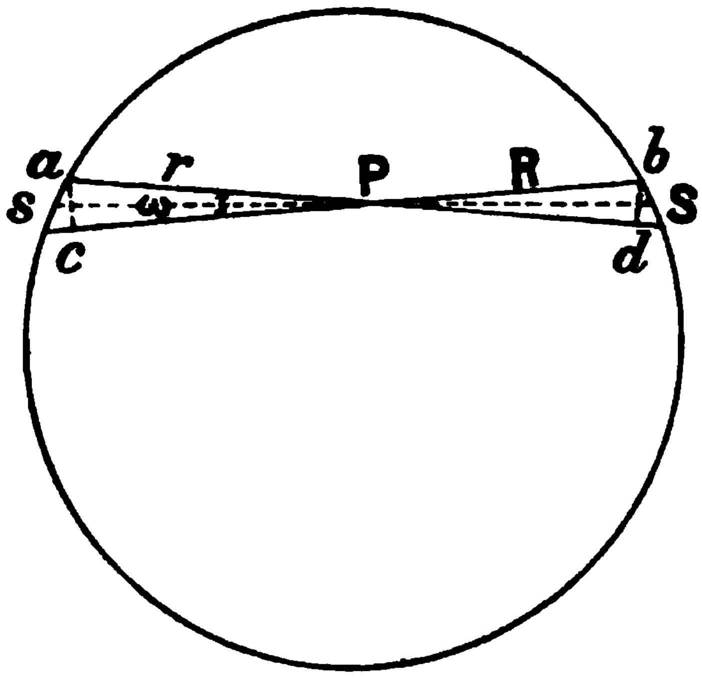


FIG. 439.

sphere can be divided up into pairs of cones of equal angle and having a common vertex at P, and if the inverse square law holds, the forces due to the charges on the portions of the surface intercepted by each pair of cones just neutralise each other, the whole electrified surface will exert no force at the point P. Since experiment shows that there is no force exerted at P, we infer that the supposition that the inverse square law is true is correct, for if the force varied as any other power of the distance, there would be some force exerted at P.

CHAPTER V

CAPACITY—ELECTRICAL ENERGY

450. Capacity of a Conductor.—There is a constant relation between the charge of a conductor and its potential, for if the density of the charge at every point of a conductor is doubled the total charge will also be doubled, and the force exerted on a unit charge, placed anywhere in the neighbourhood of the charged conductor, will also be doubled, so that the work done in removing the unit charge from the neighbourhood of the conductor to a place of zero potential will be doubled, that is, the potential of the conductor will be doubled. This constant ratio of the charge of a conductor to its potential is called the *capacity* of the conductor. Thus if a charge Q raises the potential of a conductor to V, the capacity, C, is given by the relation $C=Q/V$. If the conductor is charged to unit potential, then $V=1$ and the capacity is numerically equal to the charge necessary to charge the conductor to unit potential. Hence we may also define the capacity of a conductor as the charge which must be communicated to it to raise its potential by one unit.

451. Condensers.—We have seen in § 447 that if an uninsulated conductor is brought near a charged body, the potential of this latter is diminished on account of the induced charge on the uninsulated conductor. Hence the potential of the insulated conductor produced by a given charge is less when the uninsulated conductor is near than it is when this conductor is absent; in other words, the effect of bringing the uninsulated conductor near the charged one is to increase the capacity of this latter.

We may consider the same problem in a somewhat more direct way, if we suppose that a given conductor, say a plane AB (Fig. 440), is insulated and then charged to a potential V when at a distance from all other conductors.

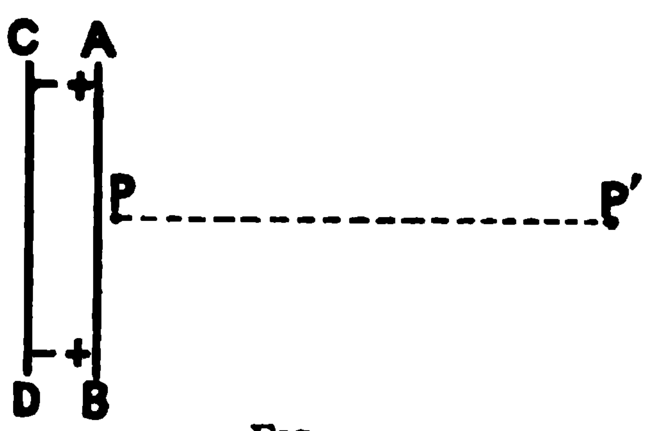


FIG. 440.

Let a second plane, which is connected with earth, be placed at such a distance from AB that its presence does not appreciably affect the electrical condition of AB. Then the work that is done in carrying a unit of positive electricity from a point P near AB to a point P′, which is at zero potential, is equal to V. Next suppose that the uninsulated plane is moved near to AB, into the position CD, so that an appreciable charge is induced on it.

The work that must now be done to move the unit charge from P to P′ by the same path as before will be less than before, for, on account of the attraction exerted on the unit charge by the negative charge induced on CD, the force exerted on the unit is everywhere less than it was before. Hence the potential of AB is less than it was before. As the plane CD is moved nearer to AB the amount of the induced negative charge increases, and the influence of this negative induced charge in diminishing the repulsive force exerted on the unit charge becomes greater and greater, and hence the potential of AB becomes less and less. The charge on AB remains however the same, and therefore, since the potential to which this charge is capable of raising AB diminishes as the uncharged and uninsulated conductor CD is brought near, it follows that the capacity of AB must increase as the conductor CB is brought near. If, instead of keeping the charge on AB constant, we had kept the potential constant, then we should have had to increase the charge on AB as the conductor CD was brought up.

An arrangement of two conductors, one of which is insulated and the other is uninsulated, placed near one another with an insulator between, is called a condenser. The name condenser was given to such an arrangement on account of the fact that the presence of the second uninsulated conductor appears to exert a condensing action on the electrical charge on the insulated conductor, so that for a given potential it can receive a much greater charge than it could without the presence of the uninsulated conductor.

The capacity of a condenser is the charge which must be communicated to the insulated conductor to raise its potential through one unit of potential. The two conductors of a condenser are sometimes called the armatures of the condenser.

The commonest form of condenser is that shown in Fig. 441, and is called a Leyden jar. It consists of a glass jar, the interior of which is coated with tinfoil up to within an inch or so of the top, and a metal knob which is in conducting communication with this inside coating. This tinfoil forms the insulated armature of the condenser, the uninsulated armature being formed by a coating of tinfoil on the outside of the jar.

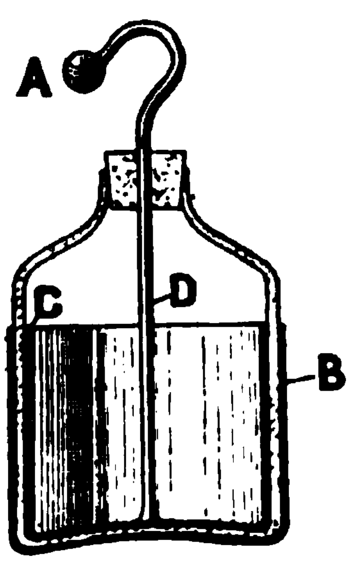


FIG. 441.

Another form of condenser which is commonly used consists of a plate of glass or some other insulating material, which is coated on each side with a sheet of tinfoil or some other conductor, a margin of an inch or so being allowed all round the edge of the glass. One coating is connected with earth, and the other forms the insulated armature of the condenser. This arrangement is sometimes called a fulminating pane.

If the insulated armature of a condenser is charged to a potential of

V, the other armature being at a potential zero, and this armature is then insulated, while the armature which was at first insulated is put to earth, this armature will not lose much of its charge, as now the rolls of the two armatures are reversed, for what was originally the induced charge is now the inducing charge, while the former inducing charge is now the induced charge. When a condenser is charged, most of the lines of force stretch across from one armature to the other, few stretching from the insulated armature to surrounding objects. Now, in order to discharge a charged conductor, the bodies on which the other ends of the tubes of force which leave the conductor terminate must be put in conducting communication with the conductor. For we may imagine that when a charged body is put to earth by means of a conducting communication, such as a wire, that the two ends of each tube of force travel along the conducting wire towards one another, the tube of force shortening up in virtue of the tension which exists along every such tube, until the two ends come together and the tube of force shrinks to nothing. In the case of the condenser, if the two armatures are put in conducting communication all the tubes of force are able to shrink to nothing, that is, the condenser becomes completely discharged. If, however, after charging the uninsulated armature is insulated, and the other armature is put in conducting communication with earth, only those tubes of force which stretch from this armature to the surrounding uninsulated conductors, such as the walls of the room, will be able to shrink and vanish. The great majority of the tubes which stretch from one armature to the other will not be able to shrink, for the armatures are not in conducting communication.

452. Specific Inductive Capacity.—If a condenser is formed by two conducting plates AB and CD (Fig. 442), placed parallel to one another, the intervening insulator being air, the capacity will have a definite value, say C. If now, while the two armatures are kept at the same distance apart, the air between the plates is replaced by some other insulator, say paraffin, the capacity of the condenser will be altered, in the case taken the capacity will be increased. We thus see that the capacity of a condenser depends not only on the geometrical conditions of the armatures, such as their size, shape, and distance apart, but also on the nature of the medium which fills the space between the plates. This fact is expressed by saying that dielectrics, as the media between the armatures are called, have different *specific inductive capacities.*

FIG. 442.

The specific inductive capacity of the air is taken as unity, and that of any other dielectric is measured by the ratio of the capacity of a condenser, of which the given substance is the dielectric, to the capacity of the same condenser when the given medium is replaced by air.

Thus if the capacity of a given condenser with air as the dielectric is C,

its capacity when the air is replaced by a dielectric of which the specific inductive capacity is K will be CK.

In order to compare the specific inductive capacities of different dielectrics, Faraday used a condenser of the form shown in Fig. 443. It consisted of an outer brass sphere, PQ, made up of two hemispheres, which fitted accurately together. This formed the uninsulated armature of the condenser, the other armature being formed by a brass sphere, C, which was held in a position concentric with the outer sphere by means of an insulating rod, A. A metal wire passing down through A allowed the inside sphere to be charged.

Two exactly similar condensers of this form were made, and one of them was charged by means of the rod B, the outside hollow sphere being connected to earth. The magnitude of the charge imparted to the condenser was then determined by touching B with a proof-plane, the charge taken away by the plane being measured with the torsion balance. The knob B was then connected with the similar knob of the other condenser, so that the two shared the charge. The charge of each was then tested by means of the proof-plane as before, and was found to be the same, thus showing that the capacity of the two condensers was the same, as from their equal size and shape ought to be the case.

FIG. 443.
(*From Ganot's "Physics."*)

Next the space, *mn*, between the inside and outside spheres in one of the condensers was filled with the medium of which the specific inductive capacity was to be determined. The other condenser was then again charged, the amount of the charge being measured as before. The knobs of the two condensers were then connected together, and the potential again measured. In the case of such a dielectric as paraffin, the potential of the two is considerably less than half the potential of the air-condenser before the two are put into communication, hence the paraffin-condenser has taken more than half the charge of the air-condenser; and since when they are connected the potential to which they are charged must be the same, it follows that the capacity of the condenser in which the dielectric is paraffin must be greater than that of the one in which the dielectric is air. In order to calculate the specific inductive capacity, K, of the paraffin, suppose that the potential to which the air-condenser was originally charged was V_1, while the potential of the two condensers when joined together is V_2. If C_1 and C_2 are the capacities of the condensers

of which the dielectrics are air and paraffin respectively, then the original charge of the air-condenser is $C_1 V_1$, while its charge after it has been put into communication with the paraffin-condenser is $V_2 C_1$. The charge of the paraffin-condenser is equal to $V_2 C_2$, but this must also be equal to the charge lost by the air-condenser, that is, to $V_1 C_1 - V_2 C_1$. Thus the specific inductive capacity of the paraffin, which by definition is equal to the ratio of the capacity of a condenser of which the dielectric is paraffin to the capacity of the same condenser when the dielectric is air, can be found. For

$$V_2 C_2 = C_1(V_1 - V_2),$$

so that

$$K = \frac{C_2}{C_1} = \frac{V_1 - V_2}{V_2}.$$

As we shall see later, the determination of the specific inductive capacity of different dielectrics is of great interest from its bearing on the electromagnetic theory of light. In the following table the values of the specific inductive capacity of some dielectrics are given. The values obtained depend, in the case of solids, on the physical condition of the solid as well as on the duration of the electrical charge employed in the measurement.

SPECIFIC INDUCTIVE CAPACITY.

Ebonite	2.5	Benzine	2.3	Air	1.0000
Glass	6.0	Ethyl alcohol	25.0	Carbon dioxide	1.0004
Shellac	3.3	Petroleum	3.1	Hydrogen	0.9997
Sulphur	3.0	Turpentine	2.2		
Mica	8.0	Vaseline	2.2		

453. Energy of a Charged Condenser.—Suppose that a condenser of capacity C is charged to a potential V, the uninsulated armature being at the potential zero. Since the potential of the one armature is V, and that of the other is O, the work done in moving a unit charge from one armature to the other will be V. Hence if we suppose that the condenser is discharged by the process of carrying the charge, one unit at a time, from one armature to the other, the work done during the transference of the first unit will be V.[1] On account of the loss of this amount of the charge the potential will be reduced to $V - 1/C$, for the original charge was VC, and the charge after the abstraction of one unit is $VC - 1$, and this charge will raise the potential of a condenser of capacity C to a potential $\dfrac{VC-1}{C}$. Hence the work done in carrying the second unit from one armature to the other will be $V - 1/C$, and so on.

[1] It will really be a little less than V, since the removal of the first unit will reduce the charge on the armature. If, however, V is great, so that the removal of a single unit makes little effect, or if we remove less than a unit each time, the error on this account can be made negligable.

Since $1/CV$ of the original charge is removed each time, after n times, where n is numerically equal to VC, the condenser will be completely discharged. Hence the total amount of work done in the discharge is the sum of the series $V+\left(V-\frac{1}{C}\right)+\left(V-\frac{2}{C}\right)+\left(V-\frac{3}{C}\right)+\ldots+\left(V-\frac{VC}{C}\right)$. This series is an arithmetical progression, and the sum of the terms is equal to half the sum of the first and last terms multiplied by the number of terms, that is, the total work done in the discharge is $\frac{1}{2}VC(V-O)$, or $\frac{1}{2}V^2C$. This expression may also be written in the forms $\frac{1}{2}QV$ and $\frac{1}{2}\frac{Q^2}{C}$, where Q is the original charge of the condenser. Since the work done in the discharge must be equal to the work done during the charge, the above expressions also express the work done in charging a condenser; in fact, these expressions give the energy of a charged condenser

FIG. 444.

(From Ganot's "Physics.")

due to the charge. We may look upon a charged condenser as possessing stored-up energy due to the strain which is set up in the dielectric, just as the coiled-up spring of a watch possesses energy due to the state of strain it is in due to its deformation.

The spark and the accompanying noise on the discharge are both evidence of the energy which is set free when a condenser is discharged. The heat produced by the discharge of a condenser may be shown and roughly measured by means of the arrangement shown in Fig. 444, and called Riess's electric thermometer. It consists of a glass globe to which is attached a narrow-bore glass tube, the end of which is connected to a reservoir containing some coloured water. A fine platinum wire is stretched across the bulb between two metal terminals which are brought

through the sides of the bulb. The pressure of the air in the bulb is adjusted so that the liquid in the tube comes to near the top, and then a charged condenser is discharged through the wire by attaching one terminal to the outside coating, and then bringing the knob of the inside coating near the other terminal. On the passage of the discharge the wire becomes heated, and the air in the thermometer expands, forcing the liquid down the tube.

If a condenser of capacity C_1 is given a charge Q, it will possess a quantity of electrical energy $Q^2/2C_1$. If now it is caused to share its charge with a second condenser of capacity C_2, by connecting together the two outside coatings, and then bringing the inside coatings into conducting communication, the charge will be shared by the two jars. The combined charge will now be equal to the original charge of the first jar, while the combined capacity is C_1+C_2. Hence the energy of the two condensers is $Q^2/2(C_1+C_2)$. Since C_1+C_2 must necessarily be greater than C_1, it follows that the energy of the two condensers is less than that of the one before it had shared its charge with the other. This loss of energy is represented by the energy spent as heat in the spark which always passes when the charge of the one jar is shared with the other. If, instead of one of the condensers being originally uncharged, they are both charged, and their inner coatings are then connected together, so that they share their charges, it can be shown that there is always less energy in the combined charges after they have been brought to the same potential than there was originally in the two separate charges, except in the case when they were originally at the same potential, when, of course, no communication of charge from one to the other takes place on their being put in communication, and thus no energy is wasted in the formation of a spark.

454. Condition of the Dielectric in an Electrical Fluid.—The important part played by the dielectric in the case of a Leyden jar may

FIG. 445.
(From Ganot's "Physics.")

be shown in a very striking manner by means of a jar such as is shown in Fig. 445, of which the armatures are removable. If such a jar is

charged, and then the inner and outer armatures are removed with a pair of insulating tongs, discharged, and then replaced, it is found that the discharge which can be obtained from the reconstructed jar is almost as strong as if the armatures had not been removed. This experiment shows how the charge of the jar is really due to some change which has been produced in the dielectric. If, after the removal of the coatings, the glass itself is discharged by passing it through a flame, no discharge can be obtained when the jar is put together again.

The phenomenon of absorption also illustrates the fact that the energy of a charged Leyden jar is stored up in the dielectric. If a jar is charged, then discharged by connecting its two coatings for a second, after a short time it will be found possible to obtain a further discharge, and after some time another, and so on, each discharge being feebler than the previous one. These residual charges, as they are called, seem to show that the electrical charge produces in the dielectric something of the nature of a sub-permanent set or strain, and that the complete recovery from this strain takes time.

The fact that the dielectric between two charged bodies is in a state of strain is shown by some experiments which are due to Kerr. We have seen, when dealing with the subject of light, that when an isotropic body is placed between crossed Nicols (§ 410) no light passes through the analysing Nicol. If, however, the isotropic body is put into a state of strain, it becomes temporarily double-refracting, and the light is able to pass through the analyser. We may, therefore, use a pair of crossed Nicols to detect the presence of a state of strain set up in an isotropic medium under any given conditions. Kerr immersed two flat metal plates, P and N (Fig. 446), in carbon bisulphide and passed through the liquid between the plates a beam of plane polarised light in a direction perpendicular to the paper. The principal plane of the polarising Nicol was parallel to AB, and on turning the analysing Nicol till its principal plane was parallel to CD, all the light was cut off. When, however, the plates P and N were brought to different potentials, so that an electric field was produced in the carbon bisulphide through which the plane polarised light was passing, the lines of force of the field being parallel to XX′, and therefore perpendicular to the direction of propagation of the light, and inclined at an angle of 45° to the principal planes of the Nicols, the light was able to pass through the second Nicol. This experiment shows that the carbon bisulphide, which under ordinary conditions is perfectly isotropic, becomes doubly refracting under the influence of the electrical field. It also shows that the supposition as to the existence of a tension along the lines of force and a pressure at right angles to the lines is justified.

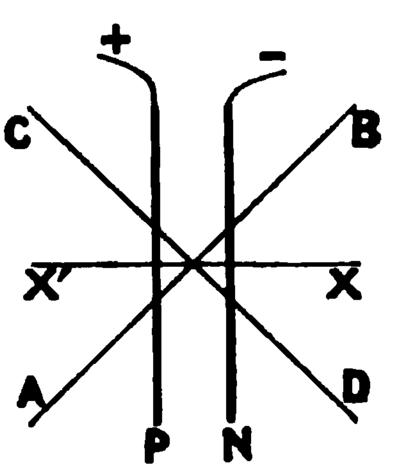


FIG. 446.

Since the dielectric plays such an important part in all electrical phenomena, it is of great importance to examine in detail the condition of the dielectric in an electrical field. In order to form a mental picture of the state of the dielectric in an electrical field, Faraday, who first drew attention to the important part played by the dielectric, imagined the system of lines and tubes of force of which we have already made some use. In the following section we shall study more in detail the property of Faraday's tubes of force, and we shall find that they very completely represent the state of an electrical field not only in a qualitative way but also quantitatively.

455. Tubes of Force.—Suppose that a very small conducting sphere A is charged with q units of positive electricity, then if A is removed from the neighbourhood of other conductors, the tubes of force will spread out from the charged sphere radially in all directions, being uniformly spaced. If an imaginary sphere were described with A as centre and a radius r (r being considerably greater than the radius of the charged sphere), the area intercepted by each tube of force on this surface would be the same. Since the charged body is charged with q units of positive electricity, q tubes of force will leave its surface, and hence q tubes of force will cut the sphere of radius r. Therefore, since the area of a sphere of radius r is $4\pi r^2$, the number of tubes of force which cut the unit area of this sphere is $q/4\pi r^2$. Now if the sphere A is sufficiently small, we may regard the charge as concentrated at a point, namely the centre, and therefore the force exerted on a unit charge at a distance r from the centre is q/r^2. But the force exerted on the unit charge is what we call the strength of the field, so that the strength of the electrical field at the surface of the sphere of radius r is q/r^2. We have just seen that the number of tubes of force per square centimetre at the surface of this sphere is $q/4\pi r^2$. Hence the strength of the field is numerically equal to the product of the number of tubes of force per square centimetre into 4π. This result may be put in a somewhat different form, for the number of tubes of force being q, and the area over which they are spread being $4\pi r^2$, the cross-section of each tube, where it cuts the sphere of radius r, is $4\pi r^2/q$. Hence the strength of the field is equal to the quotient of 4π by the cross-section, taken at right angles to the lines of force[1] of the tube of force passing through the given point. Thus if s is the cross-section of the tube of force passing through a point P, the strength of the field at P being F, we have the following relation :—

$$F = 4\pi/s.$$

If F_1 is the strength of the field at one point of a tube of force where

[1] Since the lines of force cut an equipotential surface at right angles, s is the area intercepted by the tube of force on the equipotential surface passing through the given point.

the cross-section is s_1, and F_2 is the strength of the field at another part of the tube where the cross-section is s_2, we have

$$F_1 s_1 = F_2 s_2.$$

Thus the product of the strength of the field at any point along a tube of force into the cross-section of the tube of force at the point is constant. Now the product of the electrical force in a direction at right angles to a surface into the area of that surface is called the electrical induction through the surface. Thus the induction through a normal cross-section of a tube of force is constant, for, as we have seen above, the product of the force into the area of the normal cross-section is constant, and is equal to 4π. Since the induction throughout a tube of force is constant, such a tube may be called a tube of induction. If we define a unit tube of induction as one in which the induction is unity, each of our unit tubes of force will be equal to 4π unit tubes of induction. Thus on each square centimetre of the surface of a conductor which is charged to a surface density σ there will end $4\pi\sigma$ unit tubes of induction.

456. Action of a Uniformly Charged Sphere on an External Point.—Suppose we have a conducting sphere of radius r, which is at a great distance from all other conductors and is charged with Q units of positive electricity, and we require to find the strength of the field at an external point at a distance R from the centre. Since the sphere is uniformly charged and is at a great distance from all other conductors, the lines of force must everywhere be radial, while the tubes of force will be cones having their apexes at the centre of the sphere, and will all be of the same dimensions. If a sphere were described having the same centre as the charged sphere and of radius R, each of the tubes of force would intercept the same area on this sphere, while it would cut each of the tubes at right angles. Now Q tubes leave the charged sphere, so that the area intercepted by each tube on the sphere of radius R will be $4\pi R^2/Q$. Hence if F is the strength of the field at any point on the sphere of radius R, since the force is equal to 4π divided by the normal cross-section of the tubes where they cut this sphere, we have—

$$F = 4\pi \div 4\pi R^2/Q = Q/R^2.$$

But if the whole charge, Q, of the sphere were concentrated at the centre, the force exerted at a point at a distance R would be Q/R^2. Hence the force exerted at an external point by a uniformly charged sphere is the same as would be exerted if the whole charge were concentrated at the centre of the sphere.

457. Distribution of Energy in a Field.—Suppose that we have two conducting surfaces, A and B, forming a condenser, and that they are at such a distance from all other conductors that all the lines of force which leave the one surface terminate on the other. The potential of the plate B being kept zero, let the plate A be charged with Q units

of positive electricity, so that its potential is V. Now the energy of the charged condenser is $QV/2$ and the total number of tubes of force in the field is Q, so that the quotient of the energy of the charged condenser by the number of tubes of force is $V/2$. Next suppose that the distance between the plates is decreased. The result will be that the capacity of the condenser will be increased, and so, if the charge Q on the plate A remains the same, its potential will decrease. Let us, however, increase the potential of A, that of B being still kept zero, till the potential of A has its previous value V, and let the new charge on A be Q'. The energy of the condenser is now $Q'V/2$, while Q' tubes of force occupy the field. Hence the quotient of the energy by the number of tubes is $V/2$, that is, has the same value as before. If we suppose that each tube of force contributes an equal amount to the energy of the field, then the contribution by each tube is the same in the two cases, namely $V/2$. The question, however, arises: Are we justified in supposing that each tube contributes an equal amount to the energy of the field? for some of the tubes will be short and stretch almost straight from one plate to the other, while others may be quite long and sweep round in a great curve from one plate to the other. Now we have seen above that as long as the difference of potential of the plates, that is, the difference of potential between the two ends of the tubes of force, is kept the same, the quotient obtained by dividing the energy by the number of tubes is the same whatever the relative positions of the plates. For instance we get the same result whether the plates are placed near together and parallel to one another, so that almost all the tubes stretch straight from one plate to the other, or the plates are turned so that one is at right angles to the other, and hence a large proportion of the tubes have to curve round from one plate to the other. Considerations such as these lead us to consider that each tube of force in an electrical field contributes an equal amount to the energy of the field.

We have next to see how the energy stored up in a tube of force is distributed along its length. Consider a single tube; this will start from a small area of the plate A, on which there will be a unit of positive electricity. Now if it were possible to move this portion of the surface of A along the tube of force to the plate B, this tube of force would be annihilated. In the first place let us suppose that the removal of this portion of the charge of A does not affect the potential of the plate. Under these circumstances the work done in carrying the unit from one end of the tube to the other would be V. Thus the destruction of the tube of force has been accompanied by the performance of V units of work, while the energy contained within the tube we have seen is only half this quantity. The reason for this difference is that the supposition we have made as to the potential of A remaining the same after the removal of the unit is erroneous. As the portion of the plate A carrying the unit charge is moved away from the plate A the potential will gradu-

ally fall. The result of this fall of the potential of A is that the quantity of energy stored up in each of the tubes which are left stretching from A to B is reduced. Now it can be shown that the loss on this account is exactly equal to the loss on account of the destruction of the tube, and, further, that the rates at which the losses occur as the tube is gradually destroyed by the motion of the portion of the plate A carrying the end is the same in the two cases. Hence the work done during the movement of the unit charge from a point P of the tube to a neighbouring point Q is equal to the sum of the energy contained within the portion of the tube included between P and Q, and the loss of energy of the condenser owing to the decrease of the difference of potential between its plates. Since these two losses of energy are equal, the energy contained in the part of the tube between P and Q is half the work which is done when the unit is carried from P to Q. Now the electrical force, F, varies along the tube, but if we consider the two points P and Q sufficiently near together, we may consider that F remains constant over this distance. Hence the work done in carrying the unit from P to Q will be $F.\overline{PQ}$, for the force F acts alone the tube, that is, along the direction of the path PQ. Thus the energy included in the tube of force between P and Q is $F.\overline{PQ}/2$. Hence the energy stored in unit length of the tube is $F/2$, so that the energy stored up in unit length of a tube of force is numerically equal to half the electrical force at the part of the tube considered.

The cross-section of a tube of force at a point where the force is F being $4\pi/F$, the volume of unit length of the tube in this part of the field is $4\pi/F$. But the energy stored up in unit length of the tube is $F/2$. Hence $4\pi/F$ c.c. of the field contains $F/2$ units of energy, *i.e.* ergs. Therefore the energy contained in unit volume, that is, one c.c., of the field, at a part where the electrical force is F, is $F^2/8\pi$.

If the difference of potential between the ends of a tube of force is V, and we draw the equipotential surfaces so that the difference of potential between consecutive surfaces is one unit, these surfaces will divide each tube into V small portions or cells. The whole energy stored up in the tube being $V/2$, the energy stored in each cell will be half an erg.

As we shall see later, the method of looking upon the energy possessed by a charged conductor as stored up in the field leads to many important generalisations, and the calculations we have made above will help to give the reader some mental grasp of what is implied when a certain region is said to be a field of electrical force, and to appreciate how completely this field is mapped out by the tubes of force.

458. Strength of the Field near a Charged Conductor.—If the density of the charge at a given point of the surface of a charged conductor is σ, then σ tubes of force will start from the unit of area of the surface of the conductor, and since the tubes of force intersect a conducting surface at right angles, the cross-section of a tube in the immediate neighbourhood of the surface will be $1/\sigma$. If the surface density is not

uniform, the same will still hold good in virtue of the manner in which we have defined the measure of the surface density when it is variable, for if δq is the charge on a small element of surface, δa, surrounding the given point, the surface density is $\delta q/\delta a$; but since δq tubes of force leave the small area δa, the cross-section of a single tube is $\delta a/\delta q$. Now we have seen in § 455 that the strength of a field is equal to $4\pi/s$, where s is the cross-section of the tube of force of the field at the given point. Hence the strength of the field in the immediate neighbourhood of a charged conductor, at a point where the surface density of the charge is σ, is $4\pi\sigma$. We have in the above argument assumed that the tubes of force leave the surface of the conductor at the point considered on one side only, as would be the case if the portion of the surface considered forms part of the outside of a closed surface, for under these conditions there is no force within the surface, and so all the tubes of force must leave the portion of the conductor on the one side. In the case of an unclosed conductor, such as a plane, there are two ways of regarding the problem. If, as is usual, we take σ' as the charge on both sides of unit area of the plane, then the lines of force will start out equally from each side of the plane, so that the number of tubes of force leaving each square centimetre on either side will be $\sigma'/2$, and the cross-section of a tube of force in the immediate neighbourhood of the surface will be $2/\sigma'$. Hence in this case the strength of the field will be given by $F=2\pi\sigma'$.

459. Mechanical Force Exerted on each Unit of Area of a Charged Surface.—Let ABC (Fig. 447) represent a section of a closed conductor which is charged so that the surface density over a small area of the surface at AC is σ. Let us consider the electrical force at two points, P_1 and P_2, one just outside and the other just inside the surface of the conductor. We may consider that the force at these two points is made up of two parts, namely, the force due to the portion AC of the charged surface and that due to the part CBA. Let the force at the point P_1, due to the portion CBA, be F_1, then, since the two points P_1 and P_2 are by supposition very near together, the force due to the part CBA of the conductor, which is at a comparatively great distance from both points, will be the same for both, namely F_1. Also, since the portion AC of the surface is small, it is practically plain, and so the force exerted at the two points P_1 and P_2, which are similarly situated, will be equal in magnitude but opposite in direction. Thus if the force due to AC at P_1 is F_2 acting along the outward drawn normal to the surface, the force at P_2 will be $-F_2$. Hence, adding together the two component forces for each point, we get that the force at P_1 is F_1+F_2, while the force at P_2 is F_1-F_2. But the point P_2 being within

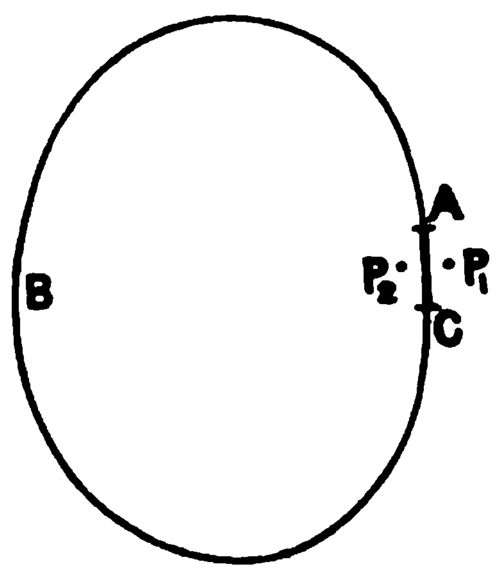


FIG. 447.

a closed conductor, the force there is zero, while we have just seen that the force at a point, such as P_1, just outside the surface of a conductor charged to a surface density σ, is $4\pi\sigma$. We therefore get that

$$F_1 + F_2 = 4\pi\sigma \text{ and } F_1 - F_2 = O.$$

$$\therefore \quad F_1 = F_2 = 2\pi\sigma.$$

Now F_1 is the force at P_1 or P_2 due to the portion CBA of the charged conductor, and therefore if we imagine that the portion AC of the surface were disconnected from the rest, without in any way altering the distribution of the charge, the electrical force acting at the point where this portion of the surface is situated is F_1 or $2\pi\sigma$. Now if a is the area of the portion AC, the charge on this portion is σa, and so the mechanical force in the direction of the normal, which this portion of the conductor would experience owing to the action of the rest of the conductor, is $\sigma a . F_1$ or $2\pi\sigma^2 a$.

Although for clearness we have supposed the portion considered to be separated from the rest of the surface, the same force is exerted, although this division does not occur. We thus see that in the case of a charged conductor, on every small element of area, a, there is exerted an outward force $2\pi\sigma^2 a$, where σ is the density of the charge on the element in question. This force acts everywhere normally to the surface, for the force at the point P_1 is normal to the surface, since the lines of force, as has been shown, everywhere cut a conducting surface at right angles. The outward normal force exerted on the unit of area is $2\pi\sigma^2$. This force is of the nature of an hydrostatic pressure, and if the conductor is expansible, as for instance is the case with a soap-bubble, it will cause the conductor to expand when it is electrified.

The electrical force, F, at P_1, that is, the force due to the whole charged conductor, being $4\pi\sigma$, we get that the mechanical force experienced by unit area of the surface of the conductor is $F\sigma/2$ or $F^2/8\pi$.

460. Tension along the Tubes of Force.—We have seen in the last section that each unit of area of a charged conductor experiences an outward mechanical force, due to the charge, which amounts per unit area to $F\sigma/2$, where F is the electrical force just outside the portion of the charged surface considered, and σ is the density of the charge on this portion of the surface. Now the number of tubes of force which start from unit area of the surface is σ, and hence, if we suppose that each of these tubes exerts a tension on the surface of the conductor equal to $F/2$, the total tension exerted on unit area will be $F\sigma/2$, that is, is of the actual amount which occurs. Hence we are able to account for the mechanical forces which act on bodies when placed in an electric field, if we suppose that each of the tubes of force is in a state of tension, the magnitude of the tension being at each point equal to half the electrical force at that point.

If we take a unit of area on an equipotential plane, passing through a point P, that is, at right angles to the tubes of force, the number of tubes which cross this unit of area will be $F/4\pi$, where F is the force at P. Hence, as the tension along each tube of force is $F/2$, the tension in the air across the unit of area is $F^2/8\pi$.

As we have already mentioned, this tension is not alone sufficient to account for the distribution of the tubes of force, but it can be shown that if in addition we imagine that there exists a pressure at right angles to the lines of force, of which the magnitude is $F^2/8\pi$ per unit area, then this distribution can be accounted for.

461. Dielectrics other than Air.—We have hitherto confined our discussion of the state of the electric field to the case where the only dielectric present was air, and we have now to proceed to consider what alterations will have to be made in the expressions we have deduced, when the whole or part of the field is occupied by other dielectrics.

Suppose that we have two infinite planes, A and B, placed parallel to one another at a distance d apart, and that the plane A is given a positive charge, such that the surface density is σ. Let the plane B be kept at zero potential, and the potential of A be V_a when the dielectric separating the planes is air, and V_k when the dielectric separating the planes has a specific inductive capacity K.

Since the planes are infinite, the field of force between the planes must be uniform, so that the tubes of force are all parallel, stretching straight across from one plane to the other, and have everywhere the same cross-section. Since the density of the charge on the plane A is σ, the number of tubes of force which leave unit area of the surface is σ, and hence the cross-section of each tube is $1/\sigma$. The cross-section of the tubes will be the same whatever the dielectric, for we suppose that the density of the charge on the plane A is kept the same in all cases.

Now the capacity of unit area of the plane A will bear the same ratio to the capacity of the whole plane as does unit area to the total area of the plane, so that we may, if we like, confine our attention to unit area taken on each of the planes. The capacity being the ratio of the charge on one plate of a condenser to the difference of potential between the plates, the capacity C_a of unit area of the planes, when the dielectric is air, is given by

$$C_a=\sigma/V_a,$$

for the charge on unit area of either plane is σ, and the difference of potential between the planes is V_a.

In the same way the capacity of unit area, when the dielectric is not air, is given by

$$C_k=\sigma/V_k.$$

Now the specific inductive capacity of the dielectric is defined as the ratio of the capacity of a condenser having the given dielectric separating

2 T

the plates to the capacity of the same condenser when the dielectric is air. Hence we have, if K is the specific inductive capacity of the dielectric,

$$K = C_k/C_a = V_a/V_k.$$

If F_a is the electrical force at any point between the planes, when the dielectric is air, the work which would have to be done to carry unit charge from one plane to the other would be $F_a d$, so that the difference of potential V_a between the planes is given by

$$V_a = F_a\, d.$$

In the same way, if F_k is the electrical force at any point between the planes, when the dielectric has a specific inductive capacity K, and V_k is the difference of potential, we have

$$V_k = F_k\, d.$$

Hence $$F_k/F_a = V_k/V_a = 1/K,$$

or $$F_k = F_a/K.$$

Now if s is the cross-section of the tubes of force, which, as we have seen, will be the same in the two cases, we have

$$F_a = 4\pi/s.$$

Hence $$F_k = 4\pi/Ks.$$

Thus in the case of a dielectric of specific inductive capacity K, the electrical force at a point is $1/K$ of the quotient of 4π by the cross-section of the tube of force at the point. Putting K equal unity, that is, dealing with air as the dielectric, we obtain the expression already found.

If N is the number of tubes which pass through unit area taken at right angles to the lines of force, we have

$$F_k = 4\pi N/K.$$

Proceeding in the same way, we may show that the energy stored up in each centimetre of a tube of force is $F_k/2$, and the energy per cubic centimetre of the dielectric is $F_k^2K/8\pi$. Also the tension in each tube is $F_k/2$, and the tension across unit area, taken at right angles to the lines of force, is $F_k^2K/8\pi$. The proof of these expressions we will, however, leave as an exercise for the reader.

462*. Force exerted between Two Small Charged Bodies when surrounded by a Dielectric other than Air.—Suppose that two small bodies, A and B, charged with Q_1 and Q_2 units of electricity respectively, are placed at a distance r apart, and are immersed in a dielectric of specific inductive capacity K. If the body A were alone present, then

the tubes of force which start from it will be uniformly distributed, and therefore since there are Q_1 tubes, the cross-section s of one of these tubes, at a distance r from the charged body, will be $4\pi r^2/Q_1$. Hence if F is the electrical force at a point P, at a distance r from the charged body, we have, as shown in the last section,

$$F=4\pi/Ks=Q_1/Kr^2.$$

Now F is the force which would be exerted on the unit charge if placed at P, and therefore if a body charged with Q_2 units were placed at P, the force exerted upon it will be FQ_2 or Q_1Q_2/Kr^2. Hence the force exerted between two small charged bodies, when immersed in a dielectric of specific inductive capacity K, is $1/K$ *th* of the force which would be exerted if they were placed at the same distance apart in air.

463*. Parallel Plate Condenser in which the Dielectric is partly Air and partly another Material.—Suppose that we have as in § 461, two infinite parallel conducting planes, A and B (Fig. 448), placed at a distance d apart, and that between them is placed an infinite parallel-faced slab, CD, of a dielectric of specific inductive capacity K, the thickness of the slab being t. Let the plane A be charged to a surface density σ, and let its potential be V_1, that of the plane B being kept at zero. The planes being infinite, the tubes of force will have everywhere the same cross-section, s, for the lines of force will all be parallel and at right angles to the planes. Thus the electrical force F_a, between A and C and between D and B, where the dielectric is air, will be given by

$$F_a=4\pi/s=4\pi\sigma,$$

while the force F_k within the slab of the dielectric will be given by

$$F_k=4\pi/Ks=4\pi\sigma/K.$$

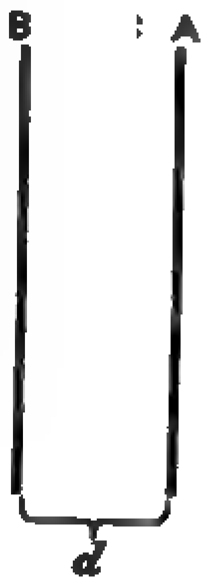


FIG. 448.

Now suppose that the unit charge is carried from near the plane A to near the plane B along a line of force. The distance moved through in air will be $d-t$, and the work done during this part of the path will be $(d-t)F_a$, or $4\pi\sigma(d-t)$. The distance moved through in the dielectric is t, and the work done during this part of the path is $F_k.t$ or $4\pi\sigma t/K$. Thus the total work done when the unit charge is carried from A to B is $4\pi\sigma\{d-t+t/K\}$. But this is equal to the difference of potential between the planes, hence

$$\begin{aligned} V_1 &= 4\pi\sigma\{d-t+t/K\} \\ &= 4\pi\sigma\{d-t(1-1/K)\}. \end{aligned}$$

Suppose now that the slab of dielectric were removed, the surface density of the charge on A remaining as before. The work which

would have to be done to carry the unit from A to B would now be $4\pi\sigma d$, that is, if V_2 is the new potential of A, that of B being still zero, we have—

$$V_2 = 4\pi\sigma d.$$

Thus the difference of potential produced by a given charge is less when the slab of dielectric is introduced, that is, of course, supposing the specific inductive capacity K of the slab is greater than that of air, and hence the capacity of the condenser is greater. The potential, and therefore also the capacity of the condenser, with air only as the dielectric will be the same as that when the slab is in place, if the distance between the plates is reduced to $d - t + t/K$, for under these circumstances the work done when the unit charge is carried from A to B will be $4\pi\sigma(d - t + t/K)$, that is, the difference of potential will have the same value as when the distance between the plates was d, but the slab was between the plates.

464*. Capacity of a Sphere when at a great Distance from all other Conductors.—Suppose that we have a sphere of radius R, surrounded by a dielectric of specific inductive capacity K, and placed at a great distance from all other conductors. Now we have seen in § 456 that the force exerted by a uniformly charged sphere at all external points is the same as would be exerted if the charge were concentrated at the centre. Hence the force exerted on the unit charge, when placed at a distance d from the centre, is Q/Kd. Suppose that we take one of the lines of force of the sphere, that is, a straight line which is the prolongation of a radius of the sphere, and that starting from a point Q_1 (Fig. 449) on this line we carry the unit charge along the line to an infinite distance from the sphere, that is, to a point where the potential is zero. The work which will have to be done to carry the unit charge from Q_1 to infinity will be equal to the potential of the point Q_1.

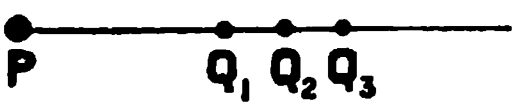


FIG. 449.

Suppose that the path is divided into a number of small elements Q_1Q_2, Q_2Q_3, &c., the distances of the points Q_1, Q_2, &c., from the centre of the sphere being d, d_1, d_2, &c. Then the force acting on the unit charge when at Q_1 is Q/Kd^2, while the force acting on it when it is at Q_2 is Q/Kd_1^2. If then the points Q_1 and Q_2 are very near together the force at Q_2 and Q_1 will be very nearly the same, and the work done while the unit charge is carried along $\overline{Q_1Q_2}$ will be very nearly equal to the product of the force at either of these points into the distance between the points. The actual amount of the work will be rather less than would be the case if the force acting was all along equal to the value it has at Q_1, and somewhat greater than if the force had everywhere the value that it has at Q_2. Thus we shall obtain a better approximation to the truth if we assume that the value of the force acting, while the

charge is moved from Q_1 to Q_2, has a value intermediate between the values it has at Q_1 and Q_2. Such an intermediate value will be obtained by taking the geometrical mean, Q/Kdd_1, of the forces at Q_1 and Q_2. Thus the work done while carrying the charge from Q_1 to Q_2 will be

$$\frac{Q}{Kdd_1}(d_1-d),$$

or

$$\frac{Q}{K}\left\{\frac{1}{d}-\frac{1}{d_1}\right\}.$$

In the same way, the work done while carrying the unit charge from Q_2 to Q_3 will be

$$\frac{Q}{K}\left\{\frac{1}{d_1}-\frac{1}{d_2}\right\},$$

and so on.

Adding together the work done over all the elements of the path we shall obtain the whole work, that is, the potential, V, of the point Q_1. Thus

$$V=\frac{Q}{K}\left\{\frac{1}{d}-\frac{1}{d_1}+\frac{1}{d_1}-\frac{1}{d_2}+\text{\&c.} \quad . \quad . \quad . \quad +\frac{1}{\infty}\right\}.$$

Now it will be seen that in this expression the distance of each of the points Q_2, Q_3, &c., occurs twice, once positively and once negatively, so that when we add together all the terms these positive and negative values will cancel, and we are left with the first term only, for the value of the last term $1/\infty$ is zero. Thus

$$V=Q/Kd.$$

Hence the potential at a point at a distance d from a uniformly charged sphere is numerically equal to the charge on the sphere divided by K times the distance of the point from the centre of the sphere. If the medium surrounding the sphere is air, the potential is obtained by putting

$$K=1.$$

If the point Q_1 is taken close to the surface of the sphere, the work which has to be done to carry the unit charge from Q_1 to infinity is the measure of the potential of the sphere. Thus the potential of a sphere, when at a great distance from all other conductors and charged with Q units, is Q/RK, or, if the medium is air, is Q/R. Now the capacity of a conductor is the ratio of the charge to the potential to which the conductor is raised by that charge. Thus the capacity of the sphere is $1/KR$, or, if surrounded by air, is $1/R$, that is, the capacity of a sphere in air is numerically equal to the radius.

465*. Capacity of a Spherical Condenser.—The problem of calculating the capacity of a system of conductors of given form is in general very difficult to solve; the case however of a condenser, such as

that shown in Fig. 443, where the two coatings are concentric spheres, can be readily obtained. Let R be the radius of the outside sphere which is connected to earth, and r the radius of the inside sphere. Let the charge on the inside sphere be Q, and the difference of potential between the two spheres be V. Then Q lines of force leave the inside sphere, and, since each of these tubes of force terminates on the inner surface of the outside sphere, there must be a charge of Q units, but of opposite sign to the charge on the inside sphere, induced on the outside sphere. If the charge on the outside sphere alone were present, the potential within this sphere would be everywhere constant, and equal to the value it has at the surface of the sphere, for, as we have shown in § 449, the potential inside a charged conductor is everywhere equal to the potential at the surface of the conductor. Hence, owing to the charge on the outside sphere, the potential everywhere inside is $-\frac{Q}{R}$, for this is the potential to which a charge $-Q$ will raise a sphere of radius R, the capacity of such a sphere being numerically equal to the radius. If the charge on the inside sphere were alone present, the potential at its surface would be Q/r. Since the potential at any point due to the simultaneous action of two charges is the sum of the potentials which each would produce if it acted alone, the potential, V, of the inside sphere, when the outer sphere is present, is given by

$$V = Q/r - Q/R = Q\left(\frac{1}{r} - \frac{1}{R}\right).$$

But the capacity, C, of the condenser is equal to Q/V. Hence

$$C = \frac{rR}{R-r}.$$

If the dielectric separating the two spheres, instead of being air, has a specific inductive capacity K, the capacity will be

$$\frac{KrR}{R-r}.$$

If the thickness, $R-r$, of the dielectric is small, the radii R and r will be very nearly equal, so that, if d is the thickness of the dielectric, $C = \frac{KR^2}{d}$. If S is the surface of the inside sphere, we have $S = 4\pi R^2$, or $R^2 = S/4\pi$. Hence under these circumstances the capacity, C, can be written $C = \frac{S}{4\pi d}$.

Although this formula only strictly applies to the case of a spherical condenser, yet it holds approximately in the case of the ordinary form of Leyden jar, in which the outside coating does not completely surround

the inside coating, and it is sometimes of use for calculating the approximate capacity of jars.

The expression for the capacity of a spherical condenser can be written in the form $C=\frac{r}{1-r/R}$. If in this expression we make R infinite, we get $C=r$, which corresponds to the case of a sphere removed from all other conductors. Hence this case may be regarded as a condenser in which the outer coating has been removed to an infinite distance. This corresponds to what was said in § 444, as to the fact that the lines of force which leave a charged body must terminate on some body, and that where they terminate will be found a charge equal in magnitude, but opposite in sign to the charge on the electrified body.

CHAPTER VI

ELECTROMETERS AND ELECTRICAL MACHINES

466. The Attracted Disc Electrometer.—Suppose that two conducting planes, AB and CED (Fig. 450), are placed parallel to one another, and at a distance d apart, so small compared to their size that the disturbing effect of their edges produces no effect at the central portions, so that the field of force between the plates is in these parts uniform. We require to find the attraction exerted on a portion, E, of the one plane of area S, when the two planes are charged to a difference of potential, V, the dielectric being air. Suppose that the surface density of the charge on AB is $+\sigma$, and that on CD is $-\sigma$, then σ tubes of force will terminate on each square centimetre of the plane CD, or at any rate on each square centimetre of the central portion, E. Hence, since these tubes are all normal to the surface of E, and each exerts a mechanical force $F/2$ (§ 460), the total attraction exerted on E by the charged plate AB is $FS\sigma/2$.

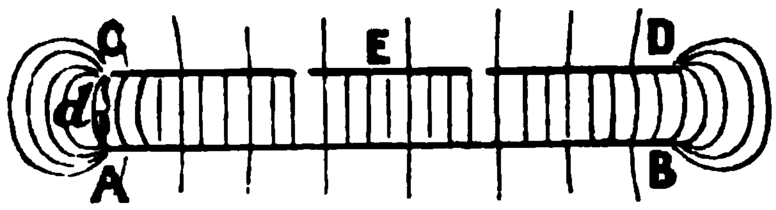


FIG. 450.

Now the cross-section of each tube of force being $1/\sigma$, the electrical force, F, acting at any point between the plates is $4\pi\sigma$. Hence the attraction, f, acting on E is given by

$$f = 2\pi S\sigma^2.$$

Since the electrical force, F, acting on the unit charge anywhere between the plates is $4\pi\sigma$, the work that must be done to carry the unit charge from one plane to the other is $4\pi\sigma d$, and therefore $V = 4\pi\sigma d$. Hence $\sigma = V/4\pi d$, and substituting this value for σ in the expression for the attraction exerted on E, we get

$$f = \frac{V^2 S}{8\pi d^2}.$$

or

$$V = \sqrt{\frac{8\pi d^2 f}{S}}.$$

Hence by measuring the force exerted on a portion of area S of the plate CD, when the distance between the plates is d, and they are charged to a difference of potential V, we can calculate the value of this

difference of potential. This then gives a method of obtaining the value of a given difference of potential in terms of the units of mass, time (involved in the value of the force), and length, that is, of determining a difference of potential in absolute units (§ 8).

The portion of the plate CD, which surrounds the part E on which the attractive force is measured, is called by Lord Kelvin, to whom the arrangement is due, the guard ring. The functions of the guard ring are simply to insure that the electrical field at the part of the plates where the attracted part E is placed shall be uniform.

In the instrument depending on this principle invented by Lord Kelvin, and called the attracted disc electrometer, or the absolute electrometer, the part E on which the force is measured consists of a metal disc supported by three springs, so that it lies concentrically within a circular hole in the guard ring, to which it is electrically connected. The springs are so arranged that when the attracted disc is attracted with a certain force by the opposite plate, AB, it lies exactly in the plane of the guard ring, as indicated by means of two sights which are attached. The plate AB can be moved in a direction parallel to its normal by means of a micrometer screw. When using the instrument the guard ring and attracted disc are connected with earth, so that their potential is zero, and the other plate is connected with the body of which the potential is to be measured. The distance between the two plates is then altered till the disc comes into its sighted position. The force necessary to bring the disc into its sighted position is determined once for all by placing weights on it, and hence, knowing this quantity (F in the formula), and also knowing the distance between the plates from the reading of the micrometer screw, the potential can be obtained.

(From Ganot's "Physics.")

467. The Quadrant Electrometer.—The absolute electrometer, although it permits of our measuring a given potential in absolute measure, is not very sensitive, and is not at all suited for comparing two small differences of potential, or for detecting the existence of a small difference of potential. Hence another form of

electrometer, called the quadrant electrometer, has been invented by Lord Kelvin. A simple form of this instrument is shown in Fig. 451. A light dumb-bell shaped aluminium needle is suspended by means of a fine metallic wire from the inside coating of a Leyden jar, the outside coating of which is connected with earth. Immediately below the needle four quadrant-shaped metal plates are supported on insulating rods. The alternate plates, A and B (Fig. 452) and C and D, are connected together by wires. The twist of the suspending wire is so adjusted that when all four quadrants are connected together, and are therefore at the same potential, the needle hangs symmetrically with reference to the two pairs of quadrants, as shown in the figure. If now the two pairs of quadrants are disconnected, and while one set are put to earth, the other, say A and B, are connected with a body which has a positive charge, then the positive charge on these two quadrants will act on the positive charge of the needle, and cause it to turn in the clockwise direction till the couple due to the torsion of the suspending wire just balances the deflecting couple due to the action of the charged quadrants on the needle.

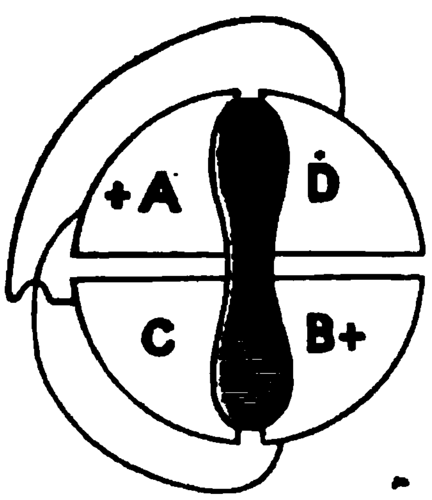


FIG. 452.

If the potential to which the needle is raised is very great compared to the difference of the potentials of the two sets of quadrants, the deflection of the needle is proportional to the difference in potential between the quadrants. Hence if the deflection produced by a known difference in potential is measured, the potential corresponding to any other deflection can be calculated. The angle through which the needle turns is generally measured by means of the motion of a spot of light reflected from a small mirror attached to the needle in the manner explained in § 332.

Since the capacity of the quadrants of a quadrant electrometer is fairly small, connecting them to a charged condenser, of which the capacity is generally enormously greater, does not appreciably alter the potential of such a condenser. If the capacity of the quadrants is of the same order as that of the body whose potential is being measured, a correction on account of the quantity of electricity taken by the quadrants must be made.

468. Electrical Machines.—We have hitherto refrained from considering the methods by which electrification, in greater quantities than can be obtained by simply rubbing a body, such as a stick of sealing-wax, can be produced, since the explanation of the manner in which the more efficient machines act involves a knowledge of the laws of induction. The oldest form of electrical machine, as the apparatus for the production of electrification is called, consisted of a glass disc or cylinder against which a pad covered with silk was pressed. When the cylinder or disc was rotated,

the friction of this pad against the glass caused the glass to become positively electrified. The electrification of the glass was collected by means of a conductor to which were attached a number of points arranged so as just to graze the surface of the glass. On account of the induction of the electrified glass these points become electrified, and since the electrical density at a sharp point is great, the negative induced charge is able to escape through the air on to the glass. This negative electricity neutralises the positive electrification on the glass, and leaves the conductor to which the points are attached positively electrified. This form of machine only works with any degree of regularity when the surrounding air is quite dry, and nowadays the only forms of machine which are employed depend on the induction produced in a conductor by an electrified body.

The simplest form of induction electrical machine is the electrophorus. This instrument is shown in Fig. 453, and consists of a disc of resin or ebonite, AB, and a metal plate, CD, which is attached to an insulating handle, E, by means of which it can be raised from the disc and carried about. The disc is electrified by friction, and the plate is placed on the top. Suppose that the material of the disc is such that it becomes positively electrified on friction, so that when thus electrified we shall have a number of tubes of force stretching from the disc to the walls of the room. On account of the fact that the surface of the disc is never quite plane, when the plate is placed on the top, contact will only take place at a very few points. Thus the plate does not become appreciably electrified by conduction from the disc, for, as it is an insulator, the electrification from those parts which are not in immediate contact with the plate are not able to travel up the disc. Hence those tubes of force which, before the plate was placed over the disc, stretched from the upper surface of the disc to the walls of the room, now terminate on the lower surface of the plate, while fresh tubes start from its upper surface and stretch away to the walls. If now the plate is earthed, that is, is put in conducting communication with the walls, the tubes which start from the upper surface of the plate will be able to shorten and vanish; that is, there will now only be the tubes which, starting from the upper surface of the disc, terminate on the lower surface of the plate. If now the plate is lifted up from the disc by its insulating handle, the distribution of the tubes of force will alter. Some of the tubes will still stretch from the disc to the plate, but as the plate gets further and further from the disc the number of these tubes gets less and less. The other tubes are, on account of the repulsion they exert on one another, driven out sideways till they meet the walls, their positive ends still remaining on the disc and their negative ends still remaining on the plate. When

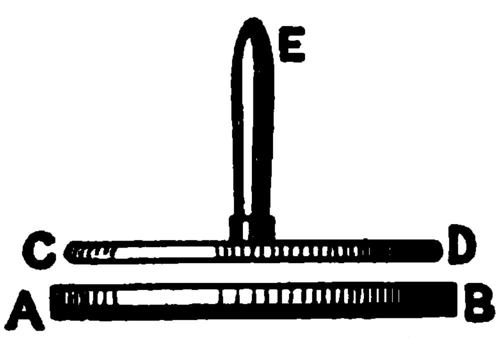


FIG. 453.

a tube meets the wall it will break in two, and we shall have two separate tubes, one starting from the disc and ending on the wall, and the other starting from the wall and ending on the plate. The tubes which still stretch from the disc to the plate correspond to that fraction of the charge of the plate which is still "bound" by the charge on the disc, while the tubes which stretch from the plate to the walls correspond to the "free" charge of the plate. When the plate is in contact with the disc and has been put to earth its potential is zero. As it is raised up from the plate, having been insulated, its potential will gradually increase; that is, in the case we have supposed, since the charge of the plate is negative, its potential will fall more and more below that of the earth.

If, after having been removed to a distance from the disc, the plate is put to earth, the tubes which start on the walls and terminate on the plate will be able to shorten and vanish, and the plate will be discharged. Now in the series of operations we have performed the charge on the disc has not been affected, and hence the plate may be replaced and the whole cycle of operations gone through again, and so on, so that the plate may be charged any number of times without recharging the disc.

It may at first sight seem as if in this way we were able to produce an indefinite amount of electricity without doing any work, and since we have seen that a charged conductor possesses energy in virtue of its charge, this would be contrary to the doctrine of the conservation of energy. It must, however, be remembered that when the plate is in contact with the disc its potential is zero after it has been put in communication with the earth, and it does not then possess any available charge. It is only after the plate has been removed from the vicinity of the inducing charge that it possesses any "free" charge. Now in order to move the plate away from the disc, work has to be done against electrical attraction between the inducing and the induced charges, or, in other words, work has to be done to stretch out the tubes of force, and it is this work which is the equivalent of the energy of the electrical charge on the plate. We are therefore here directly converting the mechanical work done by our muscles when we raise the plate into electrical energy.

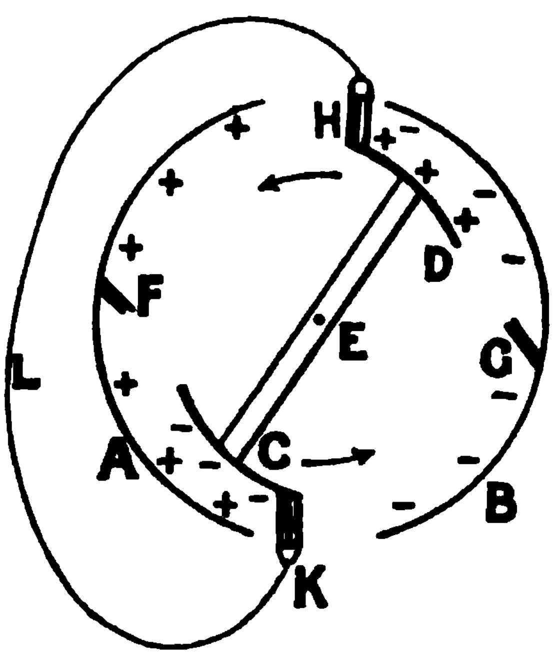


FIG. 454.

In the electrophorus a number of separate operations have to be gone through each time the plate is charged, and it naturally occurs to one to try and invent an arrangement by means of which these operations are performed automatically. The simplest of these is that due to Lord Kelvin, which is known as Thomson's Mouse Mill. It consists of two metal plates, A and B (Fig. 454), bent so as to form portions of a

cylindrical surface, these plates being carried on insulating supports. Two small metallic brushes, F and G, are attached to the inside of A and B. Two other metallic plates, C and D, are carried by an insulating arm, E, which is pivoted so that it can turn about an axis perpendicular to the plane of the figure. Lastly, there are two other metal brushes, H and K, which are in metallic connection with one another by the wire L. The brushes are so arranged that as E rotates, the plates C and D make contact with them.

Suppose that by means of an electrified rod the plate A is given an initial small positive charge, while the plate B is given a small negative charge, and that the movable arm is in the position shown in the figure. Owing to the inductive action of the charged bodies A and B, the plates C and D become electrified, one positively and the other negatively, for they form a single conductor on account of the connecting wire L. Thus we have tubes of force starting from D and ending on B. If the arm E is now rotated in the direction of the arrows, the first thing that happens is that, as D and C move round, the connection between them through the brushes H and K and the wire L is broken, while the tubes of force are drawn out. The drawn-out tubes of force will, on account of their mutual repulsion, spread out, and so most of them will come in contact with the metal plate A. Each tube, when it touches A, will divide into two parts, one part stretching from D to A, and the other from A to B, or even by further subdivision from A to the walls. As on this account as many new tubes will enter A as leave it, the charge on A will be unaltered. When D touches the brush F it becomes virtually a part of the conductor A, and thus the tubes which stretch from D to A contract to nothing; that is, the tubes which terminated on A vanish, and so on account of the new positive tubes, which were added to A when the tubes stretching from D to B split up, the charge on A is increased. In the same way the negative charge on C is transferred to B. As the rotation is continued, the plate D comes into the position in which the plate C is shown in the figure, and the whole process is repeated. Thus by the continuous rotation of the arm E carrying the two plates the charges on the conductors A and B are increased, the one being charged positively and the other negatively.

Although we have supposed an initial charge to be given to A and B, the infinitesimal charge which is induced by the friction of the movable plates on the brushes is generally sufficient to start the machine, this small charge being then increased in the manner described. If the movable arm is rotated in the opposite direction the charges on A and B are decreased, so that the arrangement is used in some instruments for adjusting the charge to a given value, for, by turning the mill in one direction or the other, the charge on a body connected to A or B can be increased or decreased at will.

The Holtz electrical machine consists of two glass discs, A and B (Fig. 455), one of which, A, is fixed, while the other, B, can be rotated by

means of a handle. The fixed plate is pierced by two apertures or windows, FF′, which are at opposite ends of the same diameter. On the face of the plate A, and on the surface turned away from the movable plate, two pieces of tinfoil, CC′, are fixed. Each of these pieces of foil has a tongue which projects through the window in the plate A. Opposite these pieces of tinfoil, and on the opposite side of the movable plate B, are placed two metal combs, DD′, which are supported on insulating stands. The machine is started by charging one of the pieces of foil, say C, with positive electricity, and rotating the glass plate in the direction of the arrows. The electricity on C acts inductively on the comb D, and as a consequence a negative change becomes spread over the surface of the glass as it is rotated. When this negative electrification reaches F′, it acts inductively on the tongue attached to the piece of tinfoil C′, and attracts an equal quantity of positive electricity to the face of the movable plate next to A. Since the thickness of the plate B is small, the positive charge derived from C′ and the negative charge derived from D, although they are on opposite sides of the plate, are so near that they practically neutralise each other's effects on any external point. Owing to the induction and to the loss of positive electricity from the tongue, the tinfoil C′ becomes negatively electrified, and this acting inductively, the comb D′ produces on the rotating glass plate a positive charge over the upper portion, which in its turn acts inductively on the tongue of C, and draws off a negative charge which neutralises the action of its own charge and increases the positive charge of C. Thus as the plate rotates the charges on C and C′ are increased, and these charges acting on the combs charge them, one positively and the other negatively, so that a body connected to one or the other may be charged either positively or negatively, or a spark may be produced between the two knobs attached to the combs. In order to increase the quantity of electricity which passes at each

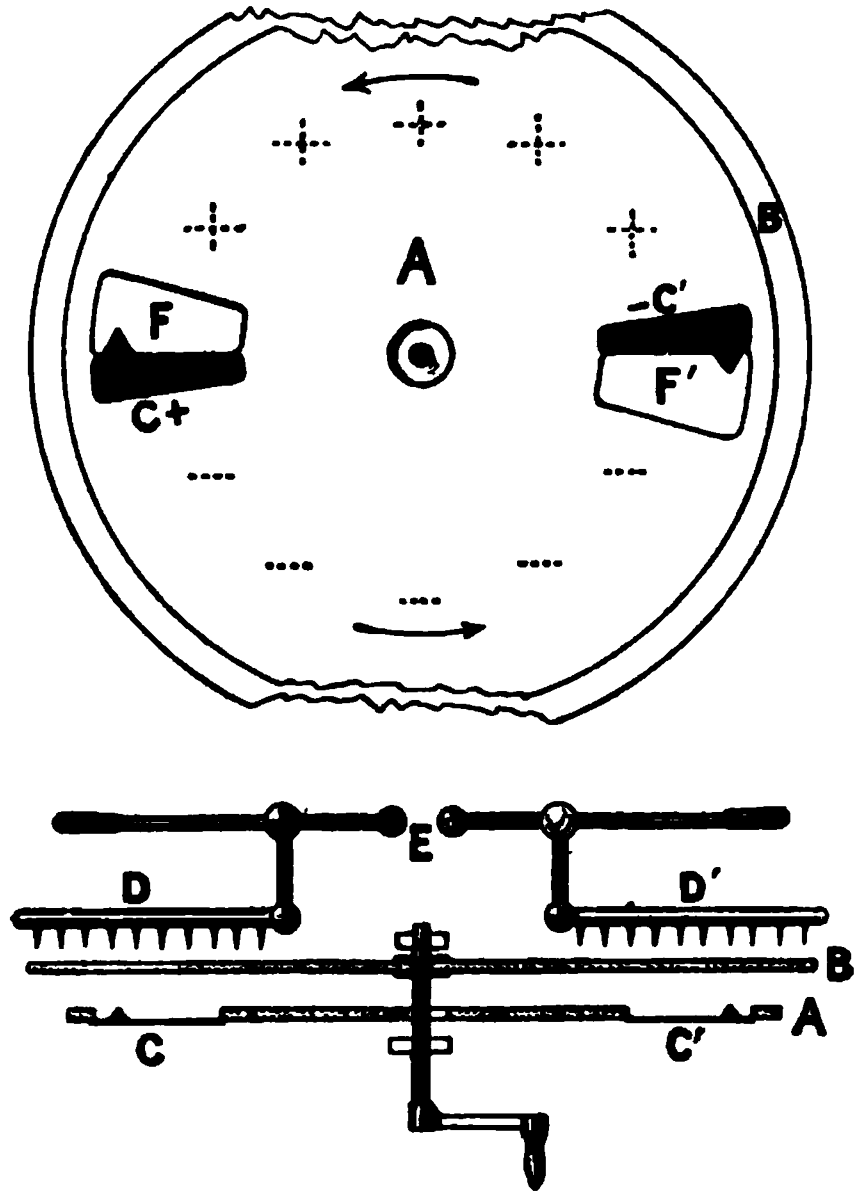


Fig. 455.

spark, the inside coatings of two small Leyden jars are connected to the combs, the outside coatings being connected together. The result is that these jars become charged, one positively and the other negatively, and when a spark passes between the knobs not only the combs but also the jars connected with them are discharged, and since the capacity of the jars is much greater than that of the combs, the quantity of electricity that passes, and hence also the brightness of the spark, is in this way much increased. It must be noticed that the addition of the jars does not produce any effect either on the quantity of electricity produced by the machine for any speed of rotation, or on the maximum difference of potential between the two combs which the machine can produce; the only effect is to store the electricity up till the potential rises to the sparking amount and to then let it discharge. Thus the sparks are less frequent than they would be without the jars, but when they do occur they are more intense.

If the combs of one machine are connected to the combs of a similar machine, and the first machine is then set in motion, the other machine will begin to turn, but in the opposite direction to that in which it is turned when it is functioning as a generator of electricity. In the first machine mechanical work is done in turning the handle, and this is converted into electrical energy, the electrical energy being transmitted to the second machine, where it is in part, at any rate, reconverted into mechanical energy.

Another form of machine by which one body can be charged to a greater potential than that of another body, although the electrification is obtained by induction from this latter, is the so-called water-dropper. This consists of an insulated metallic cylinder B (Fig. 456), down the axis of which water falls in drops from an uninsulated pipe, A. Below this cylinder is placed a second metal cylinder, C, which is fitted with a funnel of the shape shown in the figure. Some of the tubes of force which leave the charged cylinder B will terminate on the supply pipe, and of these a few will terminate on a drop of water just as it is leaving the pipe. As the drop falls it will carry one end of the tube down with it into the cylinder C, thus increasing the number of tubes which terminate on C; that is, increasing its negative charge. When the drop strikes C it forms part of this conductor, and the end of the tubes brought down by the drop will immediately travel over the surface of C to the outside, for C is very nearly a closed conductor. Thus when the drop leaves the funnel in C no tubes of force terminate on it; that is, it is uncharged. In this way the charge of C is continually increased till the leakage over the insulating support of C is equal to the rate of supply,

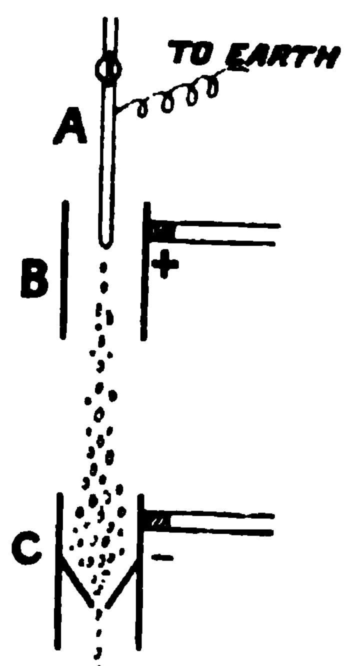


FIG. 456.

or till the cylinder is so strongly electrified that the falling drops are so strongly repelled that they no longer fall into the funnel. By having two of these arrangements, the lower cylinder of the one being connected with the upper cylinder of the other, and *vice versa*, the two will react one on the other so that the original charge on the upper cylinders will be increased, and hence also the charge produced by induction on each drop of water.

The energy necessary to produce the electrification in this apparatus is derived from the energy of the falling water. Owing to the repulsion between the charged cylinder C and each drop which carries a charge of the same sign as that on the cylinder, the resultant force acting on the drop and tending to move it downwards is less than the weight of the drop. Hence the velocity acquired by the drop in falling is less than it would be if the electrical forces were absent, so that the kinetic energy of the drop, when it strikes the funnel, is less than it otherwise would be, and the energy of the charge which it imparts to C is the equivalent of this decrease in its kinetic energy. If we suppose that the charge on each drop as it leaves the supply pipe is the same, the upward force exerted upon it by the charge on C will increase as the charge on C increases, so that the loss of kinetic energy which a drop experiences will increase as the charge on C increases. This is equivalent to an increase in the electrical energy supplied to C by the drop.

As the charge on C increases its potential will increase in the same proportion, and hence the work that must be done to bring a given charge from a place where the potential is zero to the neighbourhood of C will increase proportionally to the increase in the charge. Now the potential of the end of the pipe which is connected with earth is zero, so that each drop when it starts to fall is at a place where the potential is zero, and we thus see how the increase in the loss of kinetic energy of each drop as the potential of C increases is accounted for. This example is of interest on account of the very clear manner in which the advantage of the method we have adopted for the measure of a difference of potential is brought out, a method which may at first appear rather artificial.

PART III—ELECTRO-KINEMATICS

CHAPTER VII

THE ELECTRIC CURRENT

469. The Electric Current.—If we have two conductors at different potentials, and put them in conducting communication by means of a wire, there will be a redistribution of the electrical charges on the conductors, positive electricity leaving the conductor at the higher potential and increasing on the other. We have also seen that heat is developed in the wire, by means of which the conductors are put into communication, and we shall see that during the time that electricity is passing from one conductor to the other this conductor is the seat of many other phenomena which only last while this transference of electricity is going on. If by any means we were able to keep up the difference of potential between the two conductors, although they are connected by the wire, then this transference of electricity would continue, and the wire would continue to be the seat of a development of heat, &c. Under these circumstances the wire is said to be traversed by an electric current. The current is assumed to flow in the direction from the body at the higher potential, through the wire, to the body at the lower potential. The word current was originally used when electricity was regarded as a fluid which flowed from the conductor at the higher potential through the wire, just as a fluid flows from a place at a higher level through a pipe to a place at a lower level. As far as we are able to tell, however, the only thing that does pass is energy, this energy being in the form we call electricity, but of the nature of which we are entirely ignorant; and so far from the energy being transmitted by the wire through which the current is flowing, the accepted belief nowadays is that the energy is really transmitted by the insulating dielectric which surrounds the wire, and that the function of the wire is to direct the flow of energy. Keeping this warning in memory, it will be permissible to speak of a current of electricity flowing through a wire, and to refer to the phenomena in the space surrounding the wire as due to this current, although we no longer by these terms mean to imply any supposition as to electricity being of the nature of a fluid, or as to the wire being the path along which the energy flows.

The magnitude of the current flowing in a wire can be measured by the

quantity of electricity which passes through the wire in unit time. Thus suppose that a conductor is connected to earth by means of a fine wire, and that in order to keep it at a constant potential, V, higher than the earth, we must supply it with Q units of positive electricity per second, then the wire connecting the conductor with earth will be traversed by a current in the direction from the body to the earth, and the magnitude of this current, as measured in the units we have adopted above, is Q. If the quantity Q is constant, the wire is said to be traversed by a constant current.

470. Electromotive Force.—The cause of the electric current in the wire in the example in the above section is the fact that the two ends of the wire are at different potentials, and in such a case, where the effect of the difference of potential is to produce an electric current, that is, to move positive electrification from one place to another, it is generally spoken of as an electromotive force. Thus what we have hitherto spoken of as the difference in potential between two bodies will often, when we are dealing with electro-kinematics, be called an electromotive force between the two bodies. It must, however, be remembered that electromotive force and difference of potential are two different names for one and the same thing, and the restriction of the one term more or less rigorously to electro-statics and of the other to electro-kinematics is simply a matter of usage.

In the example given above of the wire connecting a body which was kept at a potential V to earth (the potential of the earth being taken to be zero), the electromotive force acting on the wire and to which the current is due, is V.

The electromotive force in this example might be produced by connecting the body with one of the terminals of an electrical machine, such as are described in § 468, and under these circumstances the electrical machine can be regarded as a source of electromotive force. The detailed study of the other sources of electromotive force can be undertaken with more profit at a later stage, so that for the present it will be sufficient to suppose that what is called an electric battery or voltaic cell is employed. One of the simplest of such cells is that due to Daniell, and consists of a plate of copper immersed in a solution of copper sulphate, and a plate of zinc immersed in a solution of zinc sulphate or in dilute sulphuric acid, the two solutions being separated from one another by a partition of porous earthenware. When the copper plate is connected with the zinc plate by means of a conducting wire, this wire will be traversed by a current. If the copper and zinc plates are connected with the opposite quadrants of a quadrant electrometer the needle will be deflected, showing that the copper is at the higher potential.

The consideration of the manner in which the electromotive force in this cell is developed is entirely postponed, but it may be of use to say that the energy necessary for the maintenance of the electric current in

the wire connecting the copper to the zinc is due to the chemical changes which go on in the cell when a current is passing, for the zinc is dissolved forming zinc sulphate, while the copper sulphate solution is decomposed, the copper being deposited. We have seen in § 228 that in every chemical reaction there is a definite quantity of heat absorbed or evolved, and in this case more heat would be evolved in the conversion of a given quantity of zinc into zinc sulphate than is absorbed by the splitting up of the quantity of copper sulphate solution which occurs in the same time, and it is from this surplus energy that the energy necessary for the maintenance of the electric current is derived.

471. Oersted's Experiment.—Hitherto we have not had to deal with any phenomenon connecting magnetism and electricity, although some of the points in which the two classes of phenomena resemble one another may have suggested that some connection must exist. The honour of being the first to discover any connection between electricity and magnetism belongs to Oersted, who found that a conductor in which a current is flowing exerts an action on a neighbouring magnetic needle. If a wire is stretched horizontally in the magnetic meridian, so as to be vertically over a pivoted magnetic needle, then Oersted found that the needle is deflected if a current is passed through the wire, and tends to set itself at right angles to the wire. On reversing the direction in which the current is flowing in the wire, the direction in which the north pole of the needle is deflected is also reversed. The direction of the deflection is also reversed if, instead of being placed over the needle, the wire is placed below the needle.

A number of rules have been given to remember the direction in which the needle is deflected by a conductor carrying a current in a given direction, the two most commonly employed being the following:

1. Imagine yourself swimming in the wire in the direction in which the current is flowing, and facing the magnetic needle; then the north pole will be deflected towards your left hand, the south pole being deflected in the opposite direction (Ampère's rule).

2. Place your right hand alongside the wire, with the fingers pointing in the direction in which the current is flowing, and the thumb stretched out, so that the palm of the hand is turned towards the magnet, then the north pole will be deflected towards the direction in which the thumb points.

472. Lines of Force of a Conductor conveying a Current.—If a wire through which a fairly strong current is passed is held in a vertical position, so that it passes through a hole in a horizontal plate of glass, and iron filings are scattered over the glass, on tapping the glass the filings will set themselves in curves which, as in the case of the field of a magnet, indicate the direction of the magnetic lines of force. A series of curves obtained in this way are shown in Fig. 457, and it will be seen that the lines of force consist of a series of circles, the axis of the wire being at the centre of each.

The direction in which the lines of force run can be at once obtained from either of the rules as to the direction in which a north pole is deflected in Oersted's experiment, given in the last section. Suppose that the current in the wire of Fig. 457 is running in the direction from beneath the paper to above, then, to a person swimming with the current, a north pole will be deflected to the left hand; thus the lines of force run in the anticlockwise direction. If we imagine that a corkscrew is placed in the place of the wire conveying the current, with its point in the direction in which the current is flowing, and is then turned in the direction in which it is turned in order to drive it into a cork, it will travel forward in the direction in which the current is flowing, and the direction in which it is turned will be the direction in which the lines of force of the current run. Thus the direction of the electric current, and the sense in which the lines of force run, are related to one another in the same way as are the direction of motion of an ordinary right-handed corkscrew, or other kind of screw, and the sense in which it is turned. It is of great importance, for following the forces in play between a conductor conveying a current and a magnet, or another conductor which is also conveying a current, to learn to be able at once to tell in which direction the lines of force in the neighbourhood of the conductors are running.

FIG. 457

If, instead of being straight, the conductor is bent into the form of a circle, the lines of force all thread through the space enclosed by the conducting hoop, and the general form of the lines is shown in Fig. 462.

473. Strength of the Magnetic Field due to a Current.—Since the space in the neighbourhood of a conductor in which a current is flowing is, owing to the current, a magnetic field, and that the strength of this magnetic field can be measured by the methods given in the preceding pages, we may take the strength of the field at a given distance from the conductor, which has a given shape, as a measure of the strength of the current flowing in the wire. A system of electrical units has been derived in this way, the starting-point being the strength of the magnetic field due to a conductor conveying the current. The conductor is supposed to be in the form of an arc of a circle of which the radius is one centimetre, the length of the arc being also one centimetre. Then unit current is such that the magnetic field produced at the centre of the circle, of which the conductor is an arc, is unity. Hence, since the unit magnetic field is such that the unit north pole is acted upon by a force of a dyne, the unit current may be defined in this system as such that if flowing in the arc of a circle of which the radius is one centi-

metre, the length of the arc also being one centimetre, then the force exerted on a unit pole placed at the centre of the circle will be a dyne.

We thus see that we have two ways of defining the strength of an electric current, one of them depending on the definition of the strength of an electric charge, that is, on the force exerted on one another by two charged bodies, and the other on the force which a conductor carrying a current exerts on a magnetic pole. The magnitude of this latter depends on the definition of the unit pole, which is derived from the force with which two magnetic poles act on one another. The first of these systems, namely, that depending on the force exerted between charged bodies, is called the electro-static system of electrical units, and the other is called the electro-magnetic system. On either system a consistent series of electrical and magnetic units can be built up, and in a later chapter we shall return to the relation which these two systems bear to one another, but for the present we shall make use of the electro-magnetic system, not only on account of its greater adaptability to the subjects with which we shall be dealing, but also since it is this system that is exclusively used in practice.

The unit current, defined above, is found to be too large for practical purposes, and hence the unit ordinarily employed is a tenth of the above unit. This practical unit is called the *ampere*, while the theoretical unit of which it is a tenth part, and which belongs to the centimetre-gram-second system of absolute units, is called the *c.g.s.* unit of current.

474. Units of Quantity and of Electromotive Force on the Electro-magnetic System.—In the electro-magnetic system of units, the unit quantity of electricity is the quantity of electricity which crosses a given section of a wire in which the unit current is flowing during a second. If the current is one ampere, then the unit quantity as defined above is the practical unit of quantity and is called the *coulomb*. Thus in one second one coulomb of electricity is transported past a given point by a current of one ampere.

The electro-magnetic unit of difference of potential or of electromotive force is such that if the unit electro-magnetic quantity of electricity falls through this unit of potential, the work done is one erg. This unit being very small, for practical purposes the unit adopted is 10^8 times this unit. This practical unit of electromotive force is called a *volt*.

The electromotive force of a Daniell's cell is about 1.1 volts. In order to save writing the words electromotive force at length, we shall often use the recognised abbreviation E.M.F.

475. Strength of the Field due to a Straight Conductor in which a Current is Passing.—Suppose that AB (Fig. 458) is a wire through which a current is passing, and that PCD is the line of force of the current passing through a point P, which is at a distance r from the

wire. Then the direction of the magnetic field due to the current at the point P is tangential to the line of force, *i.e.* in the direction of the arrow.

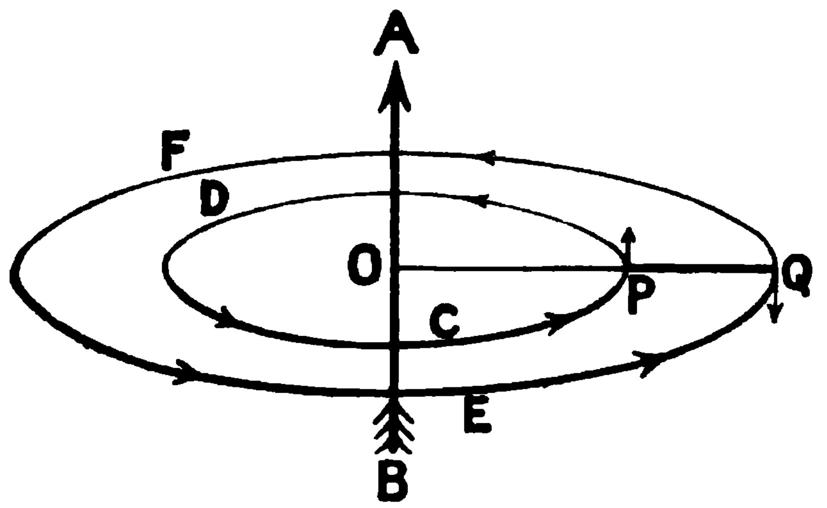


FIG. 458.

The strength of the field is obviously proportional to the strength of the current in the wire, for if a second wire, carrying an equal current, were placed alongside AB the field at P due to this current would have the same value as the field at this point due to the current in AB, and hence, when the two wires exist simultaneously, the field at P will be twice as strong as it is when only one of the wires is present. The two wires, each carrying a current C, are equivalent to a single wire in which the current is $2C$, so that doubling the current has also doubled the strength of the field at P. The rate at which the strength of the field varies with the distance r from the wire can be at once deduced from the experimental fact that if a magnet be suspended so that its two poles lie on a line passing through the wire and perpendicular to the wire, then the magnet experiences no force tending to make it rotate round the wire. Suppose that the north pole of the magnet is at P, at a distance r from the wire, while the south pole is at Q, at a distance R. Let the strength of the field at P be H, and that at Q be H', then the force exerted on the north pole P is mH, where m is the strength of the pole. The moment of this force about the axis of the wire, tending to rotate the magnet in the anticlockwise direction, is mHr. In the same way, the moment of the force due to the field of the current on the south pole at Q is $mH'R$, and tends to rotate the magnet in the clockwise direction. Since experiment shows that the magnet experiences no resultant force tending to make it rotate as a whole round the wire, these two moments must be equal and opposite. Hence

$$mHr = mH'R,$$

or

$$H/H' = R/r.$$

Hence it follows that the strength of the field in the neighbourhood of a long straight conductor conveying a current varies inversely as the distance from the conductor.

We may therefore say that the strength of the field in the neighbourhood of such a long straight wire, in which a current C is flowing, is directly proportional to the strength of the current, and inversely proportional to the distance of the point considered from the wire, or

$$H = KC/r,$$

where K is a constant. By making measurements of the strength of the field by any of the methods previously given, say by noting the deflection of a small magnetic needle when at a given distance from the wire, we can obtain the ratio of the strength of the field due to the current to the strength of the earth's field, the value of which is known, it can be shown that, if C is measured in *c.g.s.* units, the value of the constant K is 2. Hence the strength of the field at a distance r from the wire is given by $H=2C/r$, if C is measured in terms of the *c.g.s.* unit of current, and is equal to $2A/10r$ or $A/5r$, where A is the value of the current measured in amperes.

Since the strength of the magnetic field, at a distance r from a long straight wire in which a current of C *c.g.s.* units is flowing, is $2C/r$, and the circumference of a circle of radius r is $2\pi r$, the work done in carrying a unit pole round the wire along the circle of radius r will be $2\pi r \times 2C/r$ or $4\pi C$. Thus the work done in carrying a pole round the conductor is independent of the radius of the circle along which the pole is carried, and hence the work done will be the same whatever the path traversed by the pole, so long as it passes completely round the wire and then returns to its starting-point. That this is so is immediately evident, for we may split up any given path into a number of small elements which are alternately parallel to the lines of force and at right angles to the lines of force, and the work done along the sum of the portions in the direction of the lines of force will, by what has been already said, be equal to $4\pi C$, while no work will be done in the short paths which are at right angles to the lines of force, since there is no component of the force along these paths.

Although there is no couple tending to make a magnet as a whole revolve round an infinitely long straight conductor in which a current is flowing, owing to the fact that the force acting on the north pole is exactly balanced by the force acting on the south pole, yet by using a conductor of finite length we may eliminate the effect of the current on one pole, and thus allow the other pole to be moved round in the direction of the lines of force of the current. One arrangement by which this experiment may be carried out is shown in Fig. 459. A vertical axle, AB, is pivoted so that it may turn freely, and to this is attached a magnet, NS, which is bent in the manner shown in the figure. An annular mercury cup, C, is also attached to the middle of the axle, this mercury cup being in conducting communication with the axle. Thus a current can be passed down the axle entering at the upper pivot and leave by means of a wire, F, which dips in the mercury. In the portion AG of the axle there will thus be a current flowing, and the

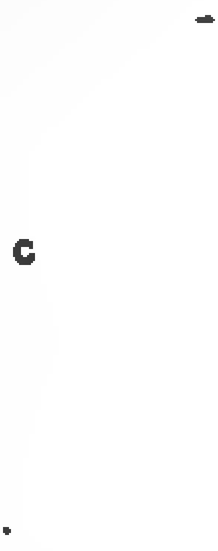


FIG. 459.

north pole of the magnet will be acted upon by this current. Since, however, there is no current in the portion GB of the axle, the south pole, S, will not be within the field of the current. Thus the magnet as a whole will be acted upon by a turning moment tending to make it rotate about the axle, the direction of motion, if we look from above, being clockwise, for the lines of force of the current from A to G, when seen from above, run in the clockwise direction.

If m is the strength of the pole of the magnet and the current flowing is C, the work done by the current on the magnet during each complete rotation is $4\pi mC$. When the steady state has been reached, that is, when the magnet is rotating with uniform velocity, this quantity of work is exactly equal to the work done against friction in moving the rotating parts of the instrument.

476. Field due to a Circular Conductor.—In the previous section we have dealt with the magnetic field of a long straight conductor in which a current is flowing, and we have now to consider the field due to conductors of other forms.

We may regard the field at any point as due to the combined action of all the small elements into which the conducting wire may be supposed to be broken up. The effect of a small element, such as we have supposed the wire to be split up into, cannot be measured experimentally, since it is impossible to obtain such an element, for the current must be conducted to and away from the element, and the magnetic effect of these conductors would have to be taken into account. Ampère, however, made a long series of experiments on the magnetic field of conductors of different forms, and he deduced from his results what would be the magnetic field of a small element of a wire in which a current C is flowing. If δs is the length of the element, and if the direction of the length of the element makes an angle θ with the line joining the centre of the element to the point where the strength of the field is to be measured, and the distance of this point from the centre of the element is r, then the strength of the field is $C\delta s . \sin \theta / r^2$.

Although the correctness of this expression cannot be directly tested by experiment, yet by its means we can calculate the strength of the field due to conductors of certain fixed forms, and then if the calculated result agrees with the value obtained experimentally, the correctness of Ampère's law will be made more and more probable as the number of experiments made is increased.

As an application of the law, we may employ it to obtain the strength of the field near a wire which is bent in the form of a circle of radius R. First let the point at which the force is to be calculated be the centre of the circle, then the angle between the element and the line joining the element to the centre of the circle is always 90°; hence $\sin \theta$ is 1 for all the elements. The distance of the point where the strength of the field is to be measured from the elements is also constant, being R the radius

of the circle. Hence the strength of the field due to each element of length δx is $C\delta x/R^2$. Further, since the direction of the lines of force due to each element are at right angles to the length of the element, the direction of the lines of force at the centre of the circle are all parallel to the axis of the circle. Thus the strength of the field due to the combined effect of all the elements is obtained by simply adding the strength due to each of them separately. The factor C/R^2 being common to all the elements, we have simply to add together the lengths of the different elements of the wire and then multiply the sum by C/R^2. But the sum of the lengths of the elements is the circumference of the circle, that is, $2\pi R$. Hence the strength of the field at the centre of the circle is $2\pi C/R$. If, instead of being a complete circle, the wire only occupies an arc of which the length is equal to the radius R of the circle, the strength of the field at the centre is C/R. If, further, the radius of the circle is one centimetre, the length of the wire also being one centimetre, the strength of the field at the centre is C. Hence if the strength of the field at the centre is unity, the current in the wire is also unity, and this result agrees with our definition of the unit current.

If instead of the wire only forming a single turn there are n turns, all of radius R, and a current C is sent through them all, since the field due to all the turns will be parallel, and the strength at the centre due to one turn is $2\pi C/R$, the strength of the field due to the n turns will be $2\pi nC/R$. If, instead of being measured in *c.g.s.* units, the current is measured in amperes, the strength of the field produced at the centre by a current of A amperes, when flowing in a circular coil of radius R and having n turns, is $2\pi nA/5R$.

To obtain the strength of the field at any point on the axis of the circle, other than the centre, we may proceed as follows. Let A and B (Fig. 460) represent the cross-section of the circular conductor by a plane (that of the paper) passing through the centre C. Let the distance of the point P from the plane of the circle be d, and the angle made by a line joining P to any point on the circumference of the circle with the axis be θ. Consider an element of the wire of length δx at A; the strength of the field due to this element at P will be $C\delta x/\overline{AP}^2$, since the line AP is at right angles to the element. Also $\overline{AP}^2 = R^2 + d^2$, so that the force due to the element is $Cdx/(R^2+d^2)$.

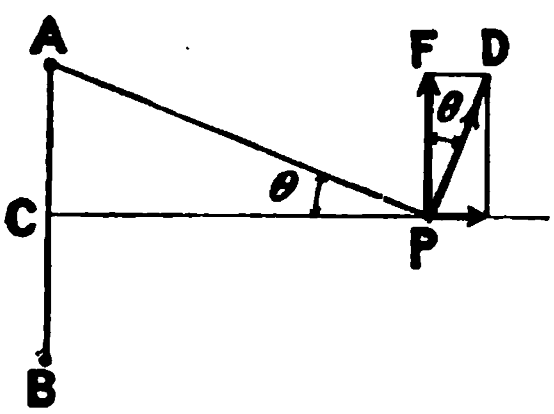


FIG. 460.

Since the lines of force of the element are circles in the plane of the paper with A as centre, the direction of the force at P is tangential to the circle in the plane passing through P perpendicular to the element described about A as centre and with AP as radius, that is, it is along PD, where PD is at right angles to AP. This force may be resolved into two

components, one along the axis of the circle and the other along the line PF at right angles to the axis. Since PD is at right angles to AP, and PF is at right angles to CP, the angle FPD is equal to θ. Hence the component of F along the axis is $\frac{C\delta x}{R^2+d^2}\sin\theta$, and the component at right angles to the axis is $\frac{C\delta x}{R^2+d^2}\cos\theta$. If we proceed in the same way for the element of length δx at B, the component along the axis will also be $\frac{C\delta x}{R^2+d^2}\sin\theta$, and the component at right angles to the axis will be $\frac{C\delta x}{R^2+d^2}\cos\theta$, but in the opposite direction to the component due to the element at A. Thus the components of the fields due to these two elements at right angles to the axis are equal and opposite, and hence are in equilibrium and neutralise each other. Since the whole circle may be split up into pairs of elements which bear to one another the same relation as do the elements at A and B, the components at right angles to the axis will on the whole neutralise one another, and in calculating the strength of the field at P we have only to consider the components parallel to the axis. The term $C\sin\theta/(R^2+d^2)$ is common to all the axial components of all the elements, and the sum of the lengths of the elements is the circumference of the circle, that is, $2\pi R$. Hence the strength of the field at P is $2\pi RC\sin\theta/(R^2+d^2)$, or, since $\sin\theta=\overline{AC}/\overline{AP}=R/\sqrt{(R+d)}$, this may be written $2\pi R^2C/(R^2+d^2)^{\frac{3}{2}}$. If there are n turns, the strength of the field at P is $2\pi nR^2C(R^2+d^2)^{\frac{3}{2}}$.

477. Galvanometers.—The deflection of a magnetic needle from its position of equilibrium in a magnetic field, either that of the earth or that due to the combined field of the earth and a magnet, by the action of the field due to a wire in which a current is flowing is the commonest way of detecting and of measuring a current. An instrument consisting essentially of a magnetic needle and a conducting wire, so arranged that when a current flows in the wire the needle is deflected and used for detecting or measuring an electric current, is called a galvanometer. Galvanometers may be divided into two great classes, namely, those used for simply detecting the passage of a current or of *comparing* the magnitude of two currents and those in which, from the magnitude of the deflection, we can calculate the magnitude of the current. In the first class the chief requisite is sensitiveness, that is, that a very small current should produce a measureable deflection of the needle ; while in the second class sensitiveness has to be subordinate to the requirement that we must be able to calculate from the dimensions of the coil, &c., the value of the field produced at the centre by unit current.

As we have seen in the last section, the strength of the field at the centre of a circular coil varies inversely as the radius of the coil, and hence if we wish to make the field produced by a given current as great

as possible, we must make the radius of the coil small. Further, since the strength of the field is directly proportional to the number of turns, the sensitiveness will obviously increase as the number of turns increases.

Suppose that the plane of the galvanometer coil is made to coincide with the magnetic meridian, so that the field due to the coil is at right angles to the magnetic meridian, then the direction in which a magnetic needle suspended at the centre of the coil will set itself will be the direction of the resultant of the field of the coil and that of the earth. If, then, by means of an auxiliary magnet the strength of the field, in which the needle hangs when no current is passing, is decreased, the resultant of this field and that due to the current in the coil will be nearer to the direction of the field due to the coil. The deflection of the needle will therefore be greater for a given strength of the coil field.

The sensitiveness of the galvanometer is also sometimes increased by employing what is called an astatic system for the needle. An astatic system consists of two magnetic needles of almost the same magnetic moment, fixed to a stem to which is attached the fibre by which they are suspended in such a way that their poles are turned in opposite directions, as shown in Fig. 461. If the magnetic moment of the magnet ns is m, while that of $n's'$ is $m+x$, where x is a small quantity, the system will set itself in the magnetic meridian with the pole n' towards the north, for the magnet $n's'$ is the stronger. If the system is deflected through an angle θ from the meridian, the couple tending to bring the needle $n's'$ back into the meridian is (§ 425) $(m+x)H \sin\theta$, while the couple tending to turn the magnet ns out of the meridian is $mH \sin\theta$. Hence the resultant couple tending to bring the system into the meridian is $xH \sin\theta$.

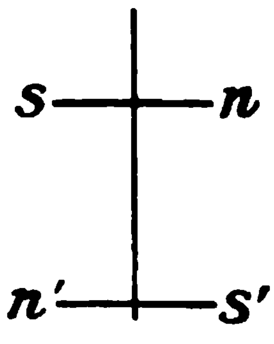


FIG. 461.

By making the quantity x small, this couple can be made as small as we like, so that if the needles are of almost the same magnetic moment, the directive couple acting on the astatic system, due to the field in which the system is suspended, is very small.

In the application of an astatic system to the galvanometer, the coil of wire is either made to surround one needle only, or two coils are employed, one round each needle, but the current is sent round the two coils in opposite directions, so that the field due to the coils in each case tends to deflect the needles in the same direction. Thus while the deflecting couple due to the field of the galvanometer coils remains the same, the directing couple which tends to bring the needles into the meridian, and hence opposes the deflection of the needle, is reduced, and the deflection produced by a given current in the coils of the galvanometer is increased.

In sensitive galvanometers the deflection of the needle is read by means of a light mirror, which is attached to the needle system, a

telescope and scale being employed as described in § 332, or the image of a slit, which is illuminated by a lamp, is thrown on a divided scale after reflection at the mirror. The image is either produced by using a concave mirror, or by placing a lens in front of the plane mirror. In both cases the angle through which the reflected beam is deflected is twice as great as the angle through which the needle is turned.

478. The Tangent Galvanometer.—In the form of galvanometer described in the last section, the field due to the coils at the place where the needles are placed is very irregular, on account of the crowding together of the turns of wire, so that as the needle turns under the influence of a current the strength of the field due to the current changes,

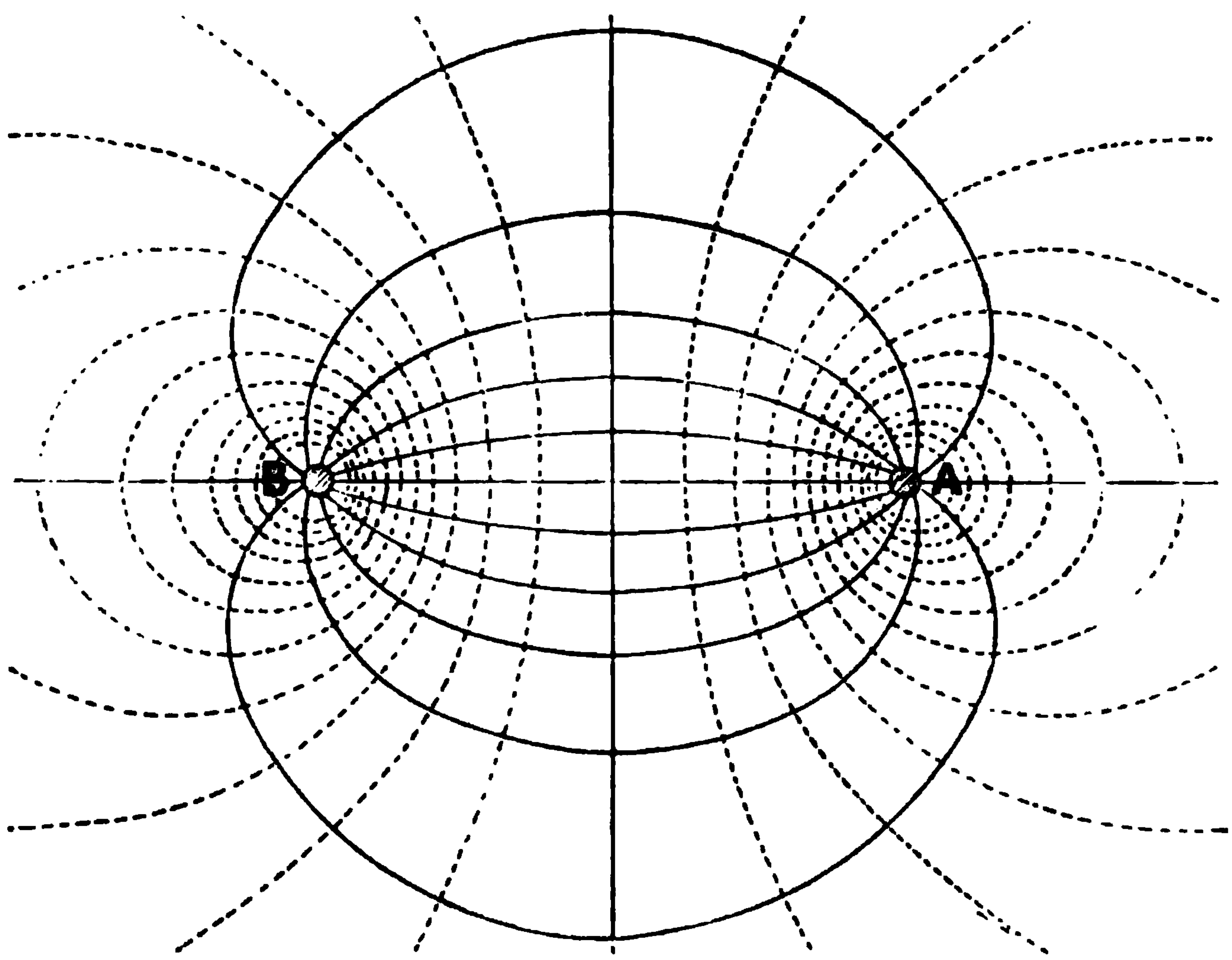


FIG. 462.

and thus, except for very small deflections, the deflection of the needle is not proportional to the strength of the current in the coils. In designing a galvanometer in which the law connecting the deflection and the strength of the current in the coils is known, it is necessary to arrange that the field due to the current at the point where the needle is placed shall be as uniform as possible.

The most common form of galvanometer for measuring currents, as distinct from detecting them, is the tangent galvanometer, so called from the fact that the currents are proportional to the tangents of the angles through which the needle is deflected. The coil of a tangent galvanometer is made of large radius, and the turns of wire are all wound in

a small groove in a metal or wooden ring, the groove being of small cross-section compared with the radius of the ring. In Fig. 462 the lines of force and the equipotential surfaces for such a coil are shown, and it will be noticed that the field is very nearly uniform near the centre, so that if the needle is small, say its length is not more than 1/20 of the diameter of the coil, the part of the field in which it hangs is practically uniform, and for any given current is of the same strength as the field at the centre of the coil. This figure also illustrates how it is that in a galvanometer, in which the length of the needle is almost as great as the diameter of the coil, the strength of the field changes as the needle is deflected.

Suppose that the coil of the tangent galvanometer is set up in the magnetic meridian, and that the value of the earth's field at the centre of the coil, where the needle is placed, is H. If there are n turns in the coil, and the mean radius of the coil is r, the cross-section of the coil being so small that the radius of each turn differs but little from the mean, the strength of the field at the centre of the coil when a current of C *c.g.s.* units is passing is $2\pi nC/r$, and the direction of the field is at right angles to the magnetic meridian.

If, under the influence of the field due to the current in the galvanometer, the needle is deflected through an angle θ from the meridian, then, as was shown in § 425, the couple tending to bring the needle back into the meridian will be $mH \sin\theta$, where m is the magnetic moment of the needle. In the same way, the couple due to the field of the coil is $2\pi nmC \cos\theta/r$. If the needle is at rest under the influence of these two couples, they must be equal, hence

$$2\pi nmC \cos\theta/r = mH \sin\theta,$$

or

$$C = \frac{rH}{2\pi n} \tan\theta.$$

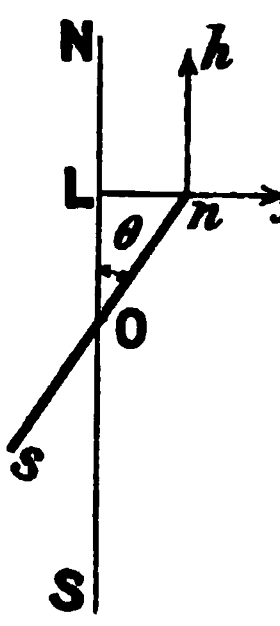


FIG. 463.

We might arrive at this result directly without using the result obtained in § 425, in the following manner. Let NS (Fig. 463) be the direction of the magnetic meridian, and hence also the trace of the plane of the coil, and let the position of the needle when deflected by the current be sOn, making an angle θ with the meridian. Then the force exerted on the pole n of the needle due to the earth's field is mH, in the direction nh parallel to SN. The turning moment of this force about O is $mH.\overline{Ln}$. But Ln is equal to $\overline{On}.\sin\theta$, or, if $2l$ is the length of the needle, to $l\sin\theta$. Hence the moment of the force due to H is $lmH\sin\theta$. In the same way, the coil has a field of strength $2\pi nC/r$ in the direction nf, and the force on the pole n, due to this field, is $2\pi mnC/r$. The moment of this force about O is $\overline{LO}.2\pi mnC/r$, or, since $\overline{LO} = l\cos\theta$,

the moment is $2\pi lmnC \cos\theta/r$. Hence, as before, the condition that the needle is in equilibrium is

$$2\pi lmnC \cos\theta/r = lmH \sin\theta,$$

or

$$C = \frac{rH}{2\pi n}\tan\theta.$$

It will be seen that for a given coil, and for a given value of the earth's field, the current passing in the coil is proportional to the tangent of the angle through which the needle is deflected. Since the value of the earth's field varies not only from place to place, but also from time

FIG. 464.

to time at the same place, it is usual to divide the expression for the current in terms of the dimensions, &c., of the coil into two parts. The quantity $r/2\pi n$, which only depends on the dimensions of the coil of the galvanometer, is called the constant of the instrument, and is generally indicated by $1/G$, so that the expression for the current which produces a deflection θ, is $C = H/G \,.\, \tan\theta$. The strengths of two currents can be compared without knowing the value of either H or G, for if θ_1 and θ_2 are the deflections produced by the currents C_1 and C_2, we have

$$C_1 = H \tan\theta_1/G \text{ and } C_2 = H \tan\theta_2/G,$$

or

$$C_1/C_2 = \tan\theta_1/\tan\theta_2.$$

The value of the constant G of the galvanometer can either be obtained by calculation from the measurements of the number of turns and of the radii made when the coil was wound, or it can be obtained experimentally by passing a current of which we know the absolute value through the coils and noting the deflection θ. Then, if the value of the earth's field H be measured in the manner given in § 432, the value of G can be calculated from the relation $G=H \tan \theta/C$.

In Helmholtz's form of tangent galvanometer the uniformity of the field near the needle is yet further insured by having two coils of equal radii placed parallel to one another at a distance apart equal to the radius of either. The needle is suspended half-way between the two coils on their common axis. The form of the lines of force for such a double coil is shown in Fig. 464, and by comparing this figure with Fig. 462 the advantage, as far as the uniformity of the field near the centre is concerned, will at once be seen.

479. The Sine Galvanometer.—If the coil of a tangent galvanometer is mounted so that it can be rotated about a vertical axis, and the angle through which it is rotated can be read off on a horizontal divided circle, another procedure for measuring a current can be employed. The coil is first turned till it lies in the magnetic meridian and the circle is read. The current is then passed, and the coil rotated about the vertical axis till the needle again lies in the plane of the coil. Using the same notation as before, and θ now indicating the angle through which the coil has been turned, the turning moment acting on the needle due to the earth's field and tending to bring it back into the meridian is $mHl \sin \theta$ as before. The moment of the force exerted by the field of the coil, which now acts at right angles to the needle, is $2\pi nmlC/r$. Hence

$$2\pi nmlC/r = mHl \sin \theta,$$

or

$$C=\frac{Hr}{2\pi n} \sin \theta = \frac{H}{G} \sin \theta.$$

Thus the current is proportional to the sine of the angle through which the coil is turned.

The usual way of performing the experiment is to turn the coil till the needle is in the plane of the coils, when the current is passing in one direction, and take the reading on the horizontal circle. The current is then reversed in direction, so that the coil has to be turned in the opposite direction. The difference between the reading of the circle when the needle is again in the plane of the coil and that obtained with the current in the opposite direction is twice the angle θ.

CHAPTER VIII

RESISTANCE

480. Ohm's Law.—If a current C is passing from a point A in a wire to another point B there must be an electromotive force between A and B, and this electromotive force is measured by the work that has to be done against electrical forces to transport the unit quantity of electricity from A to B. The connection between the electromotive force E, between any two points on the wire and the current which this E.M.F. causes in the wire, was first given by Ohm in 1827. Ohm found by experiment that the ratio of E to C was constant, so long as the physical state (temperature, &c.) of the wire between A and B was the same. This constant ratio between the electromotive force and the current is called the resistance of the conductor. Calling this quantity R, Ohm's law may be stated symbolically as follows :—

$$E/C=R,$$

or

$$C=E/R.$$

The resistance of the wire, therefore, does not depend on the strength of the current which is flowing in it. It does, however, depend on the shape and length of the wire, as also on the material of which it is composed and on the physical state of the material, such as temperature, strain, &c.

Ohm's law is entirely an empirical law, since there is no theoretical reason why it should hold. The truth of the law has, however, been subjected to most careful investigation, and it has been found that in the case of metals and electrolytes the law is true, at any rate to within one part in a hundred thousand. In the case of the passage of electricity through gases at a very low pressure it does not, however, appear to hold.

Since the resistance of a conductor is defined as the ratio of the electromotive force applied at its ends to the current passing through it, it follows that a conductor has unit resistance when unit difference of potential produces unit current in it. In the practical system the unit of resistance is called the *ohm*, and is such that the difference in potential, or the E.M.F., between the terminals of a conductor of which the

resistance is one ohm when a current of one ampere is passed through it, is one volt.

The *c.g.s.* unit of resistance is defined in the same way with reference to the *c.g.s.* units of current and E.M.F. Since the ampere is 1/10 of the *c.g.s.* unit of current and the volt is equal to 10^8 *c.g.s.* units, it follows that the ohm is equal to 10^9 *c.g.s.* units.

481. Specific Resistance.—The resistance of a given metallic conductor (the subject of the resistance of fluids is for the present postponed) depends not only on the material of which the conductor is composed, but also on the dimensions of the conductor. For a wire of a given material under constant conditions of temperature, &c., the resistance is found to be directly proportional to the length and inversely proportional to the cross-section. Hence, if l is the length and s the cross-section, the resistance R is given by

$$R = k \,.\, l/s,$$

where k is a constant depending on the nature of the material of which the wire is composed, and is called the specific resistance of the material. If both l and s are equal to unity, the resistance is equal to k. Thus we may define the specific resistance of a material as the resistance of a wire of the material of which the length is one centimetre and the cross-section is one square centimetre, or as the resistance between the opposite faces of a cube of the material of which the edge is one centimetre.

If the wire is cylindrical and of radius r, the resistance is given by $R = kl/\pi r^2$, since the cross-section is πr^2.

It is sometimes useful to deal with the reciprocal of the resistance of a conductor, and this quantity is called its conductivity. Thus if S is the conductivity of a wire, Ohm's law is expressed by $C = SE$. In the same way the specific conductivity m of a material is the reciprocal of the specific resistance, and is connected with the conductivity by the relation $S = ms/l$, the conductivity being directly proportional to the cross-section and inversely proportional to the length.

In the following table the specific resistance of some pure metals is given, but it must be remembered that a mere trace of an impurity may very largely influence the specific resistance. The specific resistance also depends to a considerable extent on the state of the material as to hardness, that is, as to whether it has been annealed or not, and if so, under what conditions the annealing has been performed.

SPECIFIC RESISTANCE.

Metal.	Resistance in Ohms of a Bar 1 cm. Long and of 1 cm.^{2} Cross-section at a Temperature of 0° C.
Silver	1.56×10^{-6}
Copper	1.3
Platinum	8.2
Iron	8.6
Tin	9.6
Zinc	5.8
Lead	19.0
Aluminium	3.2
Carbon (from Edison-Swan incandescent lamp) .	4000

482. Effect of Temperature on the Specific Resistance of Metals.—In the case of pure metals the specific resistance always in-

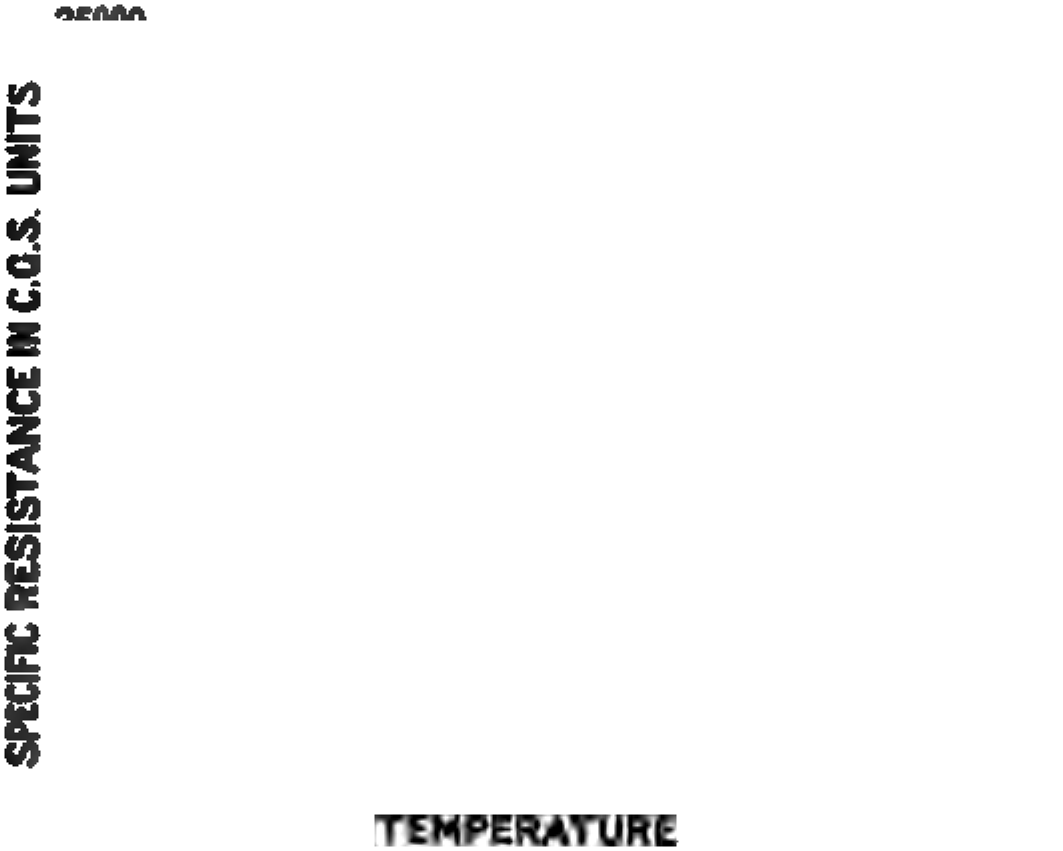


FIG. 465.

creases with increase of temperature. The change of the specific resistance with temperature of some metals, as determined by Fleming and Dewar, is shown by means of a series of curves in Fig. 465. The range of temperature employed was from about $-200°$ C. to $+200°$ C.

If any of the curves are produced backwards till they intersect the axis of X, that is, till they reach the point on the curve corresponding to zero resistance, the curious result is obtained that in almost all cases the temperature corresponding to zero resistance is $-273°$ C. Thus at the absolute zero of temperature the resistance of all pure metals would be zero.

On account of the fact that the change in the resistance of a platinum wire is often used to measure the temperature to which it is subjected, the change in resistance of this metal with temperature has been very carefully studied. As a result it has been found that if R_t is the resistance of a platinum wire at the temperature $t°$ C. on the air thermometer, and R_0 is the resistance at a temperature of 0° C., then the connection between these quantities can be expressed by an equation of the form

$$R_t/R_0 = 1 + at + bt^2.$$

In this expression a and b are constants which vary slightly from one specimen of wire to another. The value of these constants is determined by measuring the resistance of the wire when it is at three known temperatures, say in melting ice, boiling water, and boiling sulphur.

Over comparatively small ranges of temperature the increase of resistance of pure metals is very nearly proportional to the increase in temperature. Hence if R_0 is the resistance at a standard temperature, say 0° C., and R_t the resistance at a temperature t, then we may express the relation between R_0 and R_t by an expression of the form

$$R_t = R_0(1 + at).$$

The coefficient a is called the temperature coefficient of the material. For all pure metals the temperature coefficient has almost the same value, namely 0.00366. Since the coefficient of expansion of a perfect gas is 0.00366, it is seen how it is that the resistance becomes zero at the absolute zero.

483. Specific Resistance of Alloys.—As far as their electrical properties depend on those of their constituent metals, alloys may be divided into two classes. Alloys containing lead, tin, zinc, or cadmium have a specific resistance which can be calculated from that of the constituent metals if we know the proportions in which the constituent metals are present. Thus the specific resistance of an alloy containing equal masses of lead and tin will be the mean of the specific resistance of the constituent metals. In the case of most other metals the specific resistance of the alloy is much higher than would be calculated in this manner. Not only is the specific resistance of such an alloy greater than that of the constituents, but the temperature coefficient is less than that of the constituents. This is an important property of alloys from the point of view of the construction of standards of resistance, for the smaller the temperature coefficient of the material used, the less will

variations in the temperature of the standard affect its resistance. It has been found that by using manganese as a constituent of an alloy, it is possible to prepare a material of which the temperature coefficient at ordinary temperatures is either zero or even negative; that is, the resistance of the material decreases with rise of temperature.

484. Standards of Resistance.—Since, as we shall see later, the resistance of two conductors can be compared with a very high degree of accuracy, whenever possible all electrical measurements are reduced to the measurement of a resistance; and the manufacture of material standards of which the resistance is known, and of such a material that the resistance does not alter with time, is of considerable importance. Where great accuracy is desired, the forms of standard shown in Fig. 466

(*a*) FIG. 466. (*b*)

are used. The form shown at (*a*) is that adopted by the Committee of the British Association, and consists of a coil of insulated wire embedded in paraffin wax and protected by a brass case. The ends of the wire are attached to two thick copper rods, by means of which the standard can be connected with any piece of measuring apparatus, the ends of these rods being amalgamated and dipping into small mercury-cups. The alloy used for the wire is composed of two parts of silver to one part of platinum. The constancy of this alloy seems all that can be desired, but it has the disadvantage of a somewhat high temperature coefficient, and is very expensive. The form of coil shown at (*b*) is that adopted by the German National Physical Laboratory, and consists of a wire which is coated with silk and shellac, and wound on the outside of a brass cylinder, the ends of the wire being connected to stout copper connecting-rods. The wire is composed of an alloy of copper, nickel, and manganese, called manganine, which has a very small temperature coefficient. This

fact, combined with the fact that the wire not being buried in a mass of paraffin can easily take up the temperature of a liquid bath in which the coil is placed, renders this form of coil less liable to uncertainties as to the true temperature, and hence the resistance of the wire, than is the English form.

For measurements in which the greatest attainable accuracy is not desired, sets of coils are employed which are so arranged that by removing a metal plug the resistance corresponding to any coil is brought into the circuit.

485. Resistance of Systems of Conductors.—If any number of conductors are arranged so that the current goes through them all one after the other (under these circumstances the conductors are said to be arranged in series), the resistance of the arrangement is equal to the sum of the resistances of the conductors separately. That this is so follows at once from Ohm's law, for the potential of the end of one conductor is the potential at the beginning of the next, and if E_1, E_2, E_3, &c., are the electromotive forces between the ends of the conductors when a current C is passed through them all, then $E_1=R_1C$, $E_2=R_2C$, $E_3=R_3C$, &c., where R_1, R_2, R_3, . . . are the resistances of the conductors. But $E_1+E_2+E_3+\ldots=E$, the difference of potential between the two ends of the compound conductor. Hence, C being common to all the conductors, $E/C=R_1+R_2+R_3+\ldots$ But E/C is the resistance of the combined conductors, and hence the resistance is the sum of the separate resistances of the conductors.

Next let us consider the case of two conductors of resistance r_1 and r_2, which are joined together at each end. Conductors arranged in this way are said to be joined in parallel, or in parallel arc. Let E be the E.M.F. acting between the two points where the conductors join, and let the current in ABE (Fig. 467) be C_1 and that in ADE be C_2. Then from Ohm's law we have $C_1=E/r_1$ and $C_2=E/r_2$. The total current, C, flowing through the combined wires is $C=C_1+C_2$. Substituting the values for C_1 and C_2, we get $C=E(1/r_1+1/r_2)$.

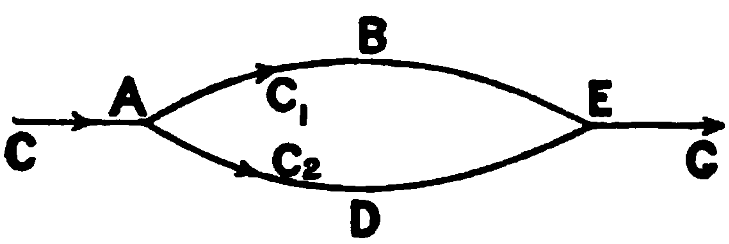


FIG 467.

Hence, as far as the current C through the combination is concerned, the two wires behave as if they were replaced by a wire of which the resistance was $\dfrac{1}{1/r_1+1/r_2}$, or $\dfrac{r_1r_2}{r_1+r_2}$. Since the reciprocal of the resistance of a wire is the conductivity, this expression amounts to a statement that the conductivity of two wires, when they are arranged in parallel, is equal to the sum of their conductivities when considered separately.

We can express the currents in the two branches in terms of the current C passing through the two wires, and the resistances of the two branches for $C_1=C-C_2=C-E/r_2$.

Hence, substituting for E in terms of C and the combined resistance, we get

$$C_1 = C - \frac{C}{r_2}\left(\frac{r_1 r_2}{r_1 + r_2}\right),$$

or

$$C_1 = \frac{Cr_2}{r_1 + r_2}.$$

In the same way

$$C_2 = \frac{Cr_1}{r_1 + r_2}.$$

If there are more than two conductors arranged in parallel, it can be shown in the same way that the conductivity of the arrangement is the sum of the conductivities of the branches separately.

486. Shunts.—If we know the resistance of each of two conductors which are arranged in parallel and measure the current, C_1, in one branch, we can calculate the current, C, which is passing through the two branches. This proposition is made use of for the purpose of measuring currents which are too large to be passed through the coils of any available galvanometer. A resistance S, called in this case a shunt, is connected in parallel with the coils of the galvanometer, of which, say, the resistance is g. Then if C is the current passing through the shunt and galvanometer, *i.e.* the current to be measured, and c is the current indicated by the galvanometer, we have, by the relation found in the last section,

$$C = \frac{S+g}{S}c.$$

Very often a galvanometer is supplied with a set of shunts of which the resistances are $g/9$, $g/99$, &c. Hence the current in the galvanometer is 1/10, 1/100, &c., of the current to be measured.

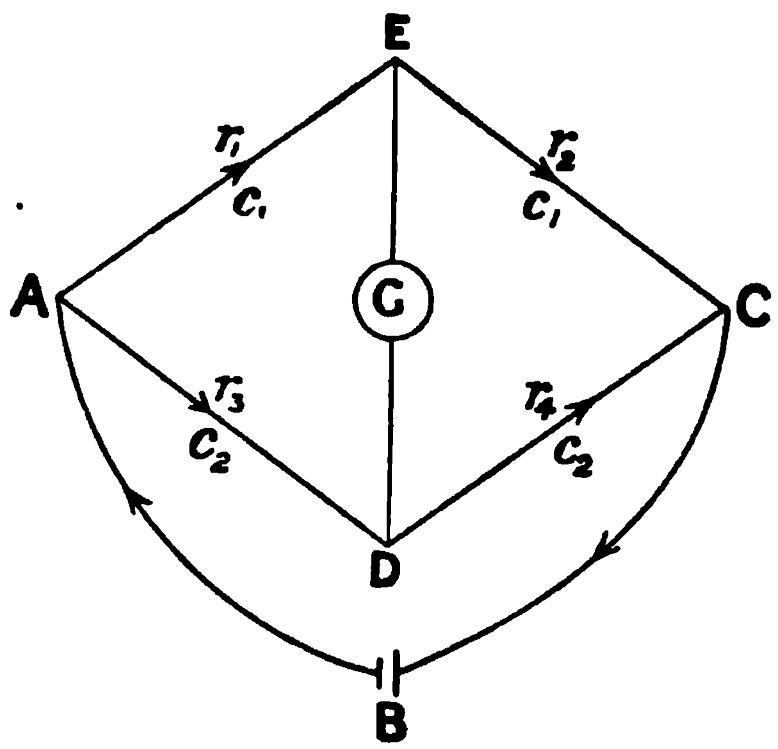


Fig. 468.

487. Wheatstone's Network of Conductors.—A system of conductors, AE, EC, AD, DC, arranged as in Fig. 468, the points A and C being connected with the poles of a battery, B, and the points D and E being connected through a galvanometer, G, is called a Wheatstone's network of conductors. If the resistances of the separate conductors are as shown on the figure, and these resistances are so adjusted that no current passes through the galvanometer, then the following relation holds :—

$$r_1/r_2 = r_3/r_4.$$

Since there is no current through the galvanometer, the potential of the points E and D must be the same. If c_1 is the current in the branch AEC, and c_2 is that in the branch ADC, we have from Ohm's law that, if e_1 is the E.M.F. between A and E, and so on for the other conductors,

$$e_1 = c_1 r_1,\ e_2 = c_1 r_2,\ e_3 = c_2 r_3,\ e_4 = c_2 r_4.$$

Since the two conductors AD and AE are in contact at A, the potential at this point must be the same for both, and we have seen that when no current is passing through the galvanometer the points E and D are at the same potential; hence the difference of potential between A and E must be the same as that between A and D or $e_1 = e_3$. In the same way $e_2 = e_4$. Hence, substituting for e_1, e_2, e_3, e_4, in terms of the currents and the resistances, we get

$$r_1 c_1 = r_3 c_2 \text{ and } r_2 c_1 = r_4 c_2.$$

Dividing one of these equations by the other,

$$r_1/r_2 = r_3/r_4,$$

or

$$r_1 r_4 = r_2 r_3.$$

This may be written $r_1 = r_2 . \dfrac{r_3}{r_4}$, which shows that if we know the value of the resistance r_2, and also the *ratio* of the resistances r_3 and r_4, we can calculate the value of the resistance r_1. Thus if we know the value of one resistance and the ratio of two others, which, when arranged together with an unknown resistance so as to form a Wheatstone's net, give no current in the galvanometer, we can immediately calculate the value of the unknown resistance.

488. Wheatstone's Bridge.—An arrangement of resistances to facilitate the measurement of a resistance by an application of the results obtained in the last section is called a Wheatstone's bridge. The simplest form of Wheatstone's bridge is shown in Fig. 469, and is called the slide-wire bridge, or sometimes, since the wire is often made a metre long, the metre bridge. In this arrangement the resistances r_3 and r_4, the ratio of whose resistances is required, are formed by the two portions of a uniform wire, AC, which is stretched alongside a divided scale. The galvanometer contact, which corresponds to the point D in Fig. 468, is formed by a sliding contact, D, which can be moved to different parts of the wire, and thus the ratio of r_3 to r_4 can be altered at will. Since the wire is uniform, the ratio of r_3 to r_4 is the same as the ratio of the lengths of the wire on the two sides of the sliding contact D. The ends of the wire are soldered to two thick copper strips F and H. The battery used to supply the current is connected to two binding-screws, B_1 and B_2, on these strips. The other terminal of the galvanometer is attached to another copper strip, I. The resistance to be measured is connected between the two binding-screws,

P_1 and P_2, or the terminals, if it is of the form shown in Fig. 466, dip into small mercury-cups at these points. The standard resistance, r_2 in the formula $r_1 = r_2 . \frac{r_3}{r_4}$, is connected with the binding-screws Q_1 and Q_2 by thick copper strips or dips into mercury-cups attached to I and H. In performing an experiment, the resistances r_1 and r_2 being arranged as described, the slider D is moved along the wire till the galvanometer is undeflected on pressing the key, by means of which the slider is put into conducting communication with the wire. If the length of the wire on the side next the unknown resistance be a, and that on the side next the

FIG. 469.

standard be b, and the resistance of this latter be R, then the resistance r of the unknown resistance will be given by

$$r = Ra/b.$$

Another form of Wheatstone's bridge, known as the Box or Post-Office form of bridge, has no stretched wire. In this form the ratio of the resistances r_3 and r_4 is not capable of being given any value we please, but the bridge is supplied with a number of coils by means of which certain fixed ratios can be obtained, the usual ratios being 1:1, 1:10, 1:100, 1:1000, 1000:1, 100:1, 10:1. In addition to these ratio coils, there are a set of coils by means of which the resistance in the arm *EC* can be made any whole number of ohms between 1 and 10,000. If the ratio of the proportional arms is 1:1, then the resistance unplugged in the third arm will be equal to the resistance being measured. If the

ratio of the proportional arms is 1 : 10, then the resistance being measured is ten times the resistance unplugged in the third arm; while if the ratio is 10 : 1, then the resistance is one-tenth, and so on.

489. The Platinum Thermometer.—For measuring temperatures much above 300° the mercury thermometer is quite unsuited, and although the air thermometer can be employed, yet its use is accompanied by so many experimental difficulties as to render it only suited for standardising other more handy forms of thermometer. The fact that the resistance of a conductor can with comparative ease be measured with a very high degree of accuracy, renders a thermometer which depends on the change of resistance of a metal wire with temperature particularly handy and accurate; the only requisite being to find a material which the resistance at any given temperature does not change with time. Further, if the temperature coefficient is fairly large, and the material will stand a high temperature without change, so much the better. Callendar and Griffiths have found that, if suitable precautions are taken as to the material on which the wire is wound, and it is properly protected from the action of certain gases, platinum fulfils these conditions thoroughly well.

The form of platinum thermometer which they have devised is shown in Fig. 470. It consists of a wire of pure platinum wound on a thin mica frame and enclosed in a glass, or, if it is required for measuring high temperatures, in a porcelain tube. The ends of the wire are connected to two thick platinum leads, P_1, P_2, by welding the platinum together. Flexible copper wires are used to connect the platinum leads to a Wheatstone's bridge, by means of which the resistance of the wire can be measured.

Fig. 470.

Since any change of temperature would affect the resistance of the platinum and copper leads, while what we require to measure is the change of resistance of the coil of thin wire only, Callendar and Griffiths have introduced a compensating device. This consists of a second pair of leads, C_1, C_2, of exactly the same resistance as the others, but which are connected together at A. These dummy leads are connected, by means of a pair of flexible copper leads of the same resistance as the others, with the opposite arm of the Wheatstone's bridge to that in which the platinum thermometer is placed. Hence, as both sets of leads are placed close together, their temperature will always be the same, and so any change in resistance produced by a variation of the temperature of the room will affect both equally. But an equal increase in the resistance in the opposite

arms of the bridge will not affect the galvanometer, and so the arrangement will be independent of any change in the temperature of the leads.

It has been found that a platinum thermometer gives consistent results up to a temperature of about 1000° C., and it has been used with much success to measure the melting-point of metals.

490. Fall of Potential along a Wire in which a Current is Passing.—When a current C is passed through a uniform wire the fall of potential along the wire is regular, the difference in the potential of any two points on the wire being proportional to the length of wire between them. This follows at once from the fact that, since the wire is uniform, the resistance per centimetre is the same at all points, and since the current is the same throughout, it follows from Ohm's law that the drop of potential per centimetre is the same throughout. The drop of potential in a conductor, or system of conductors, conveying a current can be very clearly indicated by means of a diagram in which the distance measured from some point along the conductor is taken as the abscissa, and the potential at the point is the ordinate. Thus in Fig. 471, if $\overline{AO}$ represents the potential, measured from some arbitrary zero, at one end of a uniform wire of which the length is represented by $\overline{OM}$, and the potential at the other end is represented by BM, then, since the drop of potential is uniform, the potential at any point P will be equal to the length, PC, of the ordinate drawn through P to meet the straight line joining A and B. The whole fall of potential along the wire is represented by $\overline{AN}$, where N is the point in which the line drawn parallel to OX through B meets the axis of Y. If C is the current passing through the wire, and R is its resistance, then the fall of potential or the E.M.F. between the ends will be CR, and hence $\overline{AN}$ represents CR.

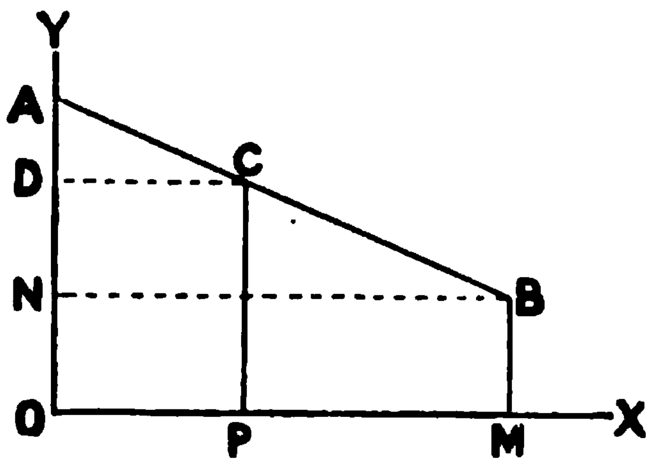


FIG. 471.

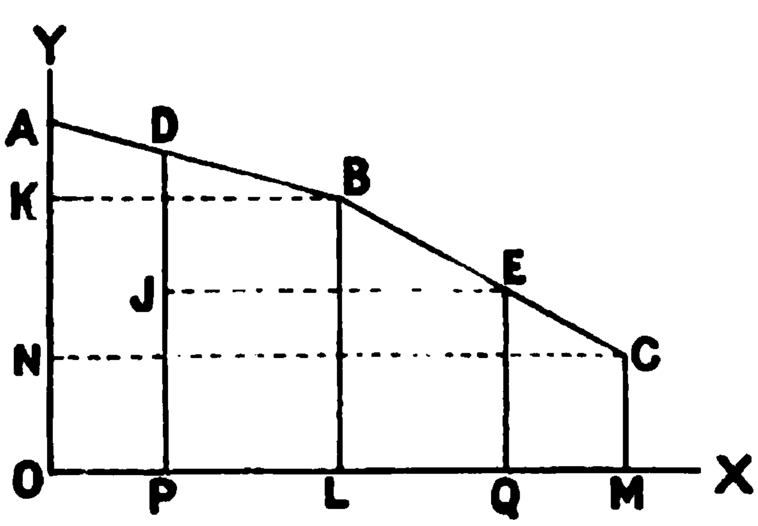


FIG. 472.

Next suppose that two wires of the same length, but one of which has a resistance per centimetre twice as great as the other, are put in series and a current C is passed through the combination. The fall of potential is shown in Fig. 472, where AB shows the fall of potential in the wire of smaller resistance, and BC that in the other wire. Since the wires are of equal length, $\overline{OL}$ is equal to $\overline{LM}$; and since the resistance of the one is twice as great as that of

the other, $\overline{KN}$ is twice $\overline{AK}$, for the fall of potential in the wire of greater resistance will be twice as great as in the other. As before, the fall of potential in each wire separately will be represented by a straight line. The E.M.F. between two points, such as P and Q, will be represented by DJ.

Returning to the case of a single uniform conductor, if two wires are connected to the conductor at the points O and P (Fig. 471), but are insulated everywhere else, then, since there is no current passing through either of these wires, the potential throughout will be same, and hence the difference of potential between the free ends will be the same as that between the points O and P, that is, will be represented by $\overline{AD}$. Suppose now we have two other wires which are connected to the two poles of some arrangement for producing a difference of potential, say a battery, and that we join the wire connected to the pole at the higher potential to the point O. Then if the difference in potential between the poles of the battery is less than the difference of potential between O and M, there will be some point on the wire through which the current is flowing which will be at the same potential as the wire connected to the other pole of the battery. Since when two points at the same potential are connected no current passes, if we connect the wire with this point P, no current will pass through the branch circuit containing the battery, the difference of potential between the points O and P being just equal to that due to the battery in the branch circuit, acting in the opposite direction. Hence if a galvanometer is included in the branch circuit, it will indicate no current when the difference of potential between O and P, due to the passage of the current *C*, is exactly equal to this E.M.F. Hence if we know the current passing in the wire OM, and the resistance of the wire between the points O and P, we can calculate the drop of potential, for it is equal to the product of the current into the resistance, and hence we have the E.M.F. of the battery in the branch circuit. This method of measuring an E.M.F. is called Poggendorff's method.

If we only wish to compare two E.M.F.'s it is no longer necessary to measure the current, for if a steady current is passed through a uniform wire, or a set of resistance coils which can be altered by a small quantity at a time, and the resistance noted which has to be intercepted between the terminals of a secondary circuit in which the E.M.F.'s and a sensitive galvanometer are included in succession, so that there is no deflection of the galvanometer, the ratio of the resistance intercepted in the two cases will, since the current is constant, be the same as the ratio of the E.M.F.'s which are to be compared. In the case where the current is passed through a uniform wire, the ratio of the lengths of wire intercepted will give the ratio of the E.M.F.'s.

491. Lines of Flow in a Conducting Sheet.—In the foregoing considerations of the flow of currents in wires we have tacitly assumed that the current was uniform over the cross-section of the conductor, and

since we have been dealing with the flow in wires, where the length of the wire was great compared with its cross-section, this was justified. We have now to consider the flow of a current in a conducting sheet, such as a sheet of tinfoil. Suppose that by means of wires attached to the foil a current is caused to flow into the sheet at a point A (Fig. 473), and to leave the sheet at the point B, so that there is a difference of potential of E between A and B. The potential will fall along the sheet from A to B, and if we draw a series of lines on the surface joining all points which are at the same potential we shall get a series of equipotential lines.

FIG. 473.

The form of these equipotential lines can be determined experimentally by using two needle-points attached by wires to the terminals of a galvanometer. One point is put in contact with the sheet at a point such as P, and the other is moved about till a point Q is found such that, when contact is made there, no current passes through the galvanometer. The fact that no current passes through the galvanometer indicates that the points P and Q are at the same potential, and are therefore on the same equipotential line. Proceeding in this way a number of points can be found, all of which are on the same equipotential line as P, and hence this line can be drawn. Then by moving the conductor from P to some other point at which the potential is different, a new equipotential line can be drawn, and so on.

In Fig. 473 the equipotential lines are shown for a sheet which is large in length and breadth compared with the distance between the points A and B, at which the current enters and leaves the sheet.

Since no current will flow from any point on a given equipotential line to any other point on this line, it is evident that the current must flow everywhere at right angles to the equipotential lines, the reasoning being exactly the same as that adopted in § 447. Hence the lines shown dotted in the figure, which everywhere cut the equipotential lines at right angles, will represent the directions in which the current will flow in the sheet.

492*. The Hall Phenomenon.—If, in the case of strip shown in Fig. 474, two conducting wires which are attached to the terminals of a sensitive galvanometer are joined to two points P and Q, which are on the same equipotential line, no current will pass through the galvanometer. If, however, the sheet is placed between the poles of a very powerful magnet, so that the lines of force of the magnetic field produced by the magnet are at right angles to the plane of the sheet, then the

galvanometer will be deflected, showing that the points P and Q are no longer equipotential. The effect of the magnetic field is to distort the lines of flow and the equipotential lines in the sheet in the way shown in Fig. 475, which corresponds to the case of a sheet of bismuth in which the lines of force of the magnetic field are passing from above the page to below, the direction of the main current in the sheet being as shown

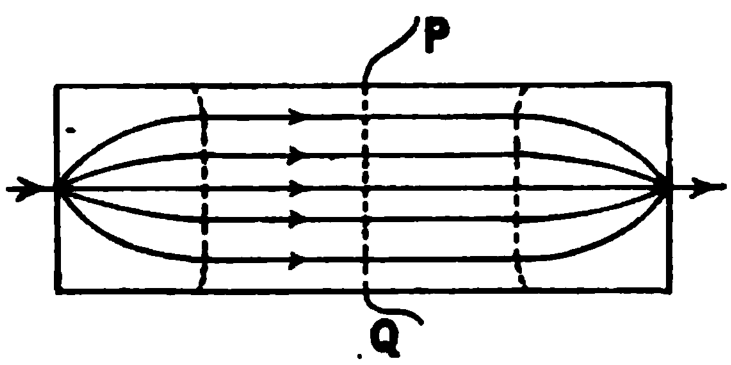


FIG. 474.

by the arrows. In the case of other metals, such as gold, the direction in which the equipotential lines and lines of flow are deflected, and hence the direction of the current in the galvanometer, is the opposite to that in the case of bismuth.

This phenomenon is referred to as Hall's phenomenon, from the name of the discoverer. We shall see later that a conductor conveying a current, when placed in a magnetic field, experiences a mechanical force tending to move it at right angles to the lines of force of the field. But

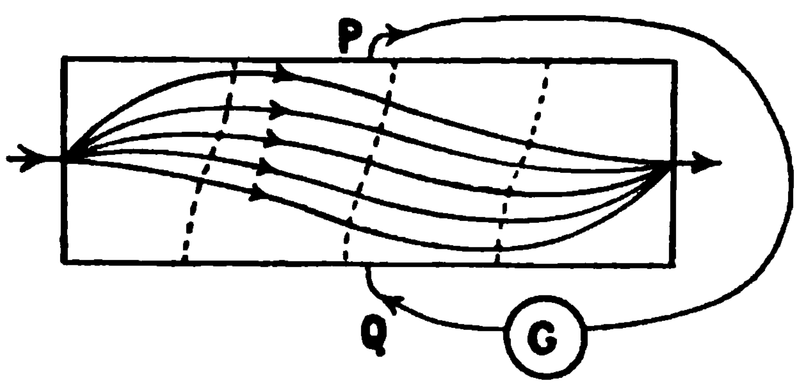


FIG. 475.

this effect on the *conductor* conveying the current must be carefully distinguished from the Hall effect, which refers to the *current* in the conductor.

The magnitude of the Hall effect is excessively small, and it is only with very strong magnetic fields and a very delicate galvanometer that it can be detected at all.

CHAPTER IX

JOULE'S LAW

493. Joule's Law.—We have defined the E.M.F. between two points, A and B, on a conductor through which a current is flowing as being equal to the work done in transporting the unit quantity of electricity from one of these points to the other. We have also seen that, if R is the resistance between the points and C is the current passing, the E.M.F. between the points is given by $E=CR$. Now a current C will transport C units of electricity past each point of the conductor in the unit of time, hence the work done in a second in driving the current from the point A to the point B is given by EC or $RC.C$, that is, RC^2. If the current continues for a time t, the work done will be represented by the equation $W=RC^2t$. The work thus done appears as heat developed in the conductor through which the current is flowing.

In order to get the amount of heat developed in the conductor in thermal units we have to divide this result by the mechanical equivalent of heat, that is, by 4.19×10^7 (§ 250). Hence the quantity of heat, H, developed in a wire of which the resistance is R *c.g.s.* units by a current of C *c.g.s.* units in t seconds is given by

$$H=\frac{RC^2t}{4.19\times 10^7}\text{ calories}$$
$$=0.2387\times 10^{-7}.RC^2t\text{ calories.}$$

If the electrical quantities are measured in the practical units, since a current of C *c.g.s.* units is equal to 10 C amperes, and a resistance of R *c.g.s.* units is equal to $10^{-9}R$ ohms, the heat developed in a conductor of which the resistance is O ohms, by a current of A amperes, in a time t seconds, will be 0.2387 OA^2t calories. The work done in one second by a current of one ampere when passing through a resistance of one ohm, that is, when passing between two points between which there exists a difference of potential of one volt, is 10^7 ergs, or in thermal units 0.2387 calories, and is called a joule.

The law that the heat developed in a conductor is proportional to the square of the current and to the resistance was discovered experimentally by Joule in 1841, and hence is known as Joule's law.

The quantity of energy which becomes converted into heat when a given current flows through a given conductor is independent of the direc-

tion in which the current flows, for, as the current always flows from the point at the higher potential to the point at the lower potential, if we reverse the direction in which the current is flowing, this means that we have reversed the direction in which the E.M.F. is acting; and since the resistance of a conductor is independent of the direction in which the current is flowing, the conditions, as far as the work done by the current and hence the heat developed, are exactly the same when the current is reversed as they were before. Thus the passage of a current through a conductor of finite resistance is always accompanied by the conversion of a definite quantity of electrical energy into the form of heat. Since when the current is reversed the conversion into heat continues at the same rate as before this conversion of electrical energy into heat, when a current passes through a conductor, is an irreversible process. As we shall see later, there are conditions under which heat developed at a given point due to the passage of a current is a reversible process, so that on reversing the current heat is now absorbed at the point; in the case of heat developed according to Joule's law, however, this is never the case. Since in many cases the heat produced according to Joule's law is simply a waste of energy, it is important to reduce it to a minimum. This can be done, if we suppose that a given current has to be transmitted, by reducing the resistance of the conducting wires. Since, as we have seen, the resistance of a pure metal at the absolute zero appears to be zero, a current could be passed through such a conductor at the absolute zero without the production of any heat and the consequent loss of electrical energy.

494. The Mechanical Equivalent of Heat derived from Electrical Experiments.—Since the heat developed by a current of A amperes in a wire of resistance O ohms, in a time t seconds, is equal to $OA^2t \times 10^7$ ergs or $OA^2t \times 10^7/J$ calories, where J is the value of Joule's equivalent, if we measure the heat developed by means of a calorimeter, and also the current A and the resistance O of the wire, or, what comes to the same thing, the E.M.F. between the ends of the wire, we can at once calculate the value of J. A most careful determination of the value of J by this method has been carried out by Griffiths. His apparatus consisted of a coil of platinum wire through which the current could be passed, and which had two wires attached, so that the difference of potential between the ends of the coil could be compared with the E.M.F. of a standard Clark cell (§ 554), by the method given in § 490. This coil was contained inside a closed calorimeter, which was itself placed inside a large steel chamber, the space between the outside of the calorimeter and the walls of this vessel being exhausted of air so as to reduce the loss of heat due to convection. The calorimeter contained, in addition to the coil, a stirrer, which was rotated at a high speed, so as to insure the water inside being thoroughly well mixed. The temperature of the water in the calorimeter was measured with a platinum thermometer, and the resistance of the

coil at different temperatures was determined, so that, knowing the E.M.F. between the terminals and the temperature, the resistance of the coil and the rate at which heat was being developed by the current could be calculated. A certain amount of heat was also developed by the friction of the stirrer against the water. The amount of heat thus developed at different rates of stirring was determined by making observations of the rise in temperature of the calorimeter, due to the stirring alone, when no current was passing through the coil. The water value of the calorimeter and of the stirrer and coil was determined by making experiments with various quantities of water in the calorimeter.

As has been given already, the value obtained for the mechanical equivalent of heat was 4.1940×10^7 ergs per calorie. The accuracy of this value depends, of course, on the accuracy with which the values of the electrical quantities are known in terms of the fundamental units.

495. The Incandescent Electric Lamp.—The heat developed in a conductor by the passage of an electric current is made use of in the electric incandescent lamp. The modern form of lamp consists of a fine carbon filament enclosed in a glass globe from which the air has been exhausted. The resistance of a carbon filament being fairly great, the heat developed is sufficient to raise the temperature to such a point that the filament glows with a bright white light.

The energy which becomes converted from the electrical form in the filament is partly given out from the lamp in the form of light and partly as heat. The object of the lamp-maker is to produce a lamp in which the proportion of the energy used up to produce heat, and which as far as the production of light is concerned is completely wasted, is reduced to a minimum. It is found that the energy which has to be supplied to a lamp in order to produce a light of one candle-power decreases as the temperature of the filament is increased, so that from this point of view it is an advantage to run the lamps as bright as possible. It is, however, found that when the temperature is raised above a certain point the filament soon gives way, so that the life of the lamp is short. The resistance of the filaments of the lamps are adjusted so that when the E.M.F. between the ends of the filament has certain definite values, such as 100 volts or 200 volts, the current which passes, according to Ohm's law, will raise the temperature of the filament to the greatest value which will not endanger its life. With most of the incandescent lamps of good make the energy consumed to produce a light of one candle-power is about four watts, or, since one watt is equal to 10^7 ergs per second, is about 4×10^7 ergs per second.

The number of watts required per candle-power increases very rapidly as the E.M.F. between the ends of the filament is reduced below the value for which the lamp is intended, so that it is very wasteful to run the lamps at a low voltage, although by this means the life of the filament may be increased.

496. The Arc Lamp.—Another source of light which depends on the conversion of electric energy into light is the arc lamp. When two rods, composed of the carbon which is deposited inside the retorts used in the manufacture of illuminating gas, connected to two conductors which are at a difference of potential of about 60 volts, are first brought into contact and are then gradually separated for a short distance the current continues to pass, and where it crosses the air space between the carbon rods an intense light is emitted. This arrangement constitutes an electric arc, and it is found that the carbon rod which is at the higher potential, that is, from which the current goes, is eaten away more rapidly than the other carbon. If an image of the arc is projected on a screen, it is seen that the carbon which is at the higher potential, the positive carbon as it is usually called, is worn slightly hollow, and that the greater proportion of the light is emitted from this hollow, which is called the crater of the arc. The end of the negative carbon, *i.e.* that at the lower potential, becomes worn to somewhat of a point. In order to allow for the wearing away of the carbon rods, they are held in an arrangement by which they are automatically brought nearer together as the ends wear away, so that the length of the arc is maintained constant. If by chance the distance between the carbons becomes too great the current will cease to pass, and the arc cannot again be started till the carbons are brought into contact and then separated. Hence the lamp is fitted with an electrical arrangement by which, directly the current ceases, the rods are first brought together, and then, when the current passes, are again separated to the correct distance.

An ordinary arc requires about one watt for each candle-power produced, so that the energy consumed in order to produce a given quantity of light by means of an arc is much less than is required when incandescent lamps are used.

497. The Electric Furnace.—The temperature of the arc is extremely high, it having been estimated to be about 8000° C., and this high temperature has been utilised for melting refractory substances and for conducting chemical processes which require a very high temperature.

The form of electric furnace used by Moisson on his important researches at high temperatures consists of a block of lime or fireclay through which pass two thick rods of carbon, which act as electrodes for the supply of the current. The arc is formed between the ends of these rods just above the substance which is to be heated, which is contained in a small crucible placed in a cavity cut in the block of lime.

PART IV.—THERMO-ELECTRICITY

CHAPTER X

THERMO-ELECTRICITY

498. Thermo-Electric Junction.—In 1821, while making experiments on the difference of potential which appears to exist between two different metals when placed in contact (see § 545), Seebeck noticed that if a circuit is formed which is composed of two wires of different metals joined together at their ends, and if the junctions are at different temperatures, a current will in general be produced in the circuit. Thus if two copper wires, which are connected to the terminals of a galvanometer, are connected at their other ends to a piece of iron wire, and one of the junctions of the copper and iron is heated, a current will be indicated by the deflection of the galvanometer. The direction of the current will be from the hot to the cold junction in the iron. This current is said to be a thermo-electric current.

If, while the cold junction is kept at a constant temperature, the temperature of the hot junction is gradually raised, the current in the circuit will gradually increase up to a certain point, this temperature being called the neutral point for the two given metals. If the temperature of the hot junction is raised above the neutral point the current in the circuit will decrease, till, when the temperature of the hot junction is as much above the neutral point as that of the cold junction is below, there will be no current in the circuit; while if the temperature of the hot junction is yet further raised, the direction of the current will be reversed.

It is possible to arrange the metals in a series such that if wires of any two of them are joined together to form a circuit, and one of the junctions is heated, the thermo-electric current will in the first metal on the list go from the hot junction to the cold, it being supposed that the mean temperature of the hot and cold junctions has some given value. The following is such a thermo-electric series for a mean temperature of about 50° C.: antimony, iron, zinc, silver, tin, copper, bismuth. Of course the order of the metals will vary with the temperature, for the neutral temperature for some of the combinations is quite low, and the neutral points for the different combinations vary very much.

499. Thermo-Electric Power and Thermo-Electric Diagrams.—If we take some metal as a standard—lead is the one usually taken—and form a thermo-electric couple between this metal and another, and

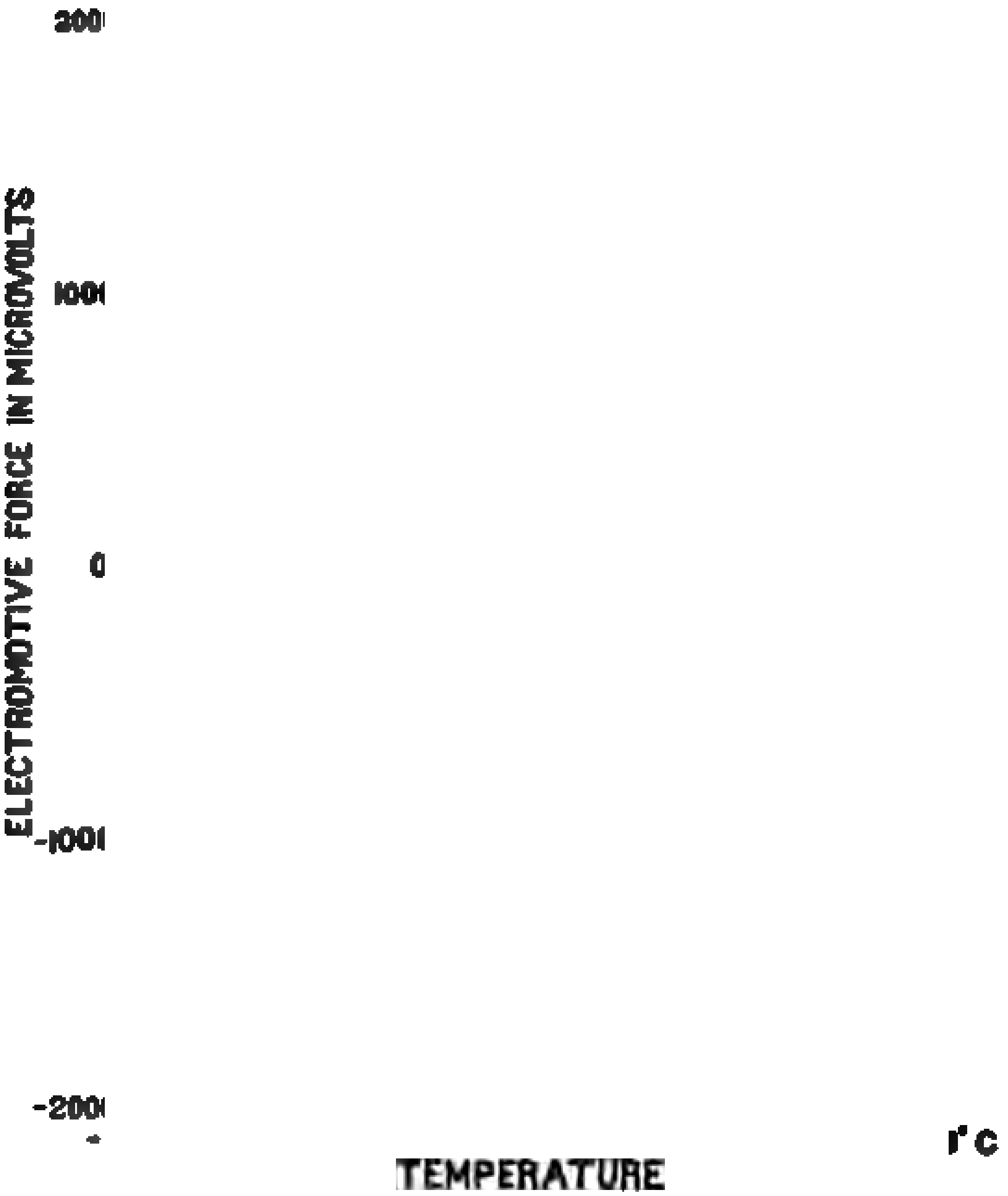


FIG. 476.

attach a lead wire to the other end of this second metal, then, if while the temperature of one of the lead-metal junctions is kept constant, say at 0° C., the temperature of the other junction is raised to different temperatures, there will be produced a difference of potential between

the free ends of the lead wires. This difference of potential is the thermo-electric E.M.F. due to the temperature difference between the hot and cold junctions. If a series of measurements of this thermo-electric E.M.F. is made for different temperatures of the hot junction, that of the cold junction being kept constant at 0° C., and the results are plotted on a curve, the temperatures of the hot junction being taken as abscissæ and the corresponding E.M.F.'s as ordinates, a curve of the form of those shown in Fig. 476 will be obtained. These curves represent the thermo-electric E.M.F.'s of some metals taken with reference to lead in terms of microvolts, *i.e.* 10^{-6} volts, and degrees Centigrade.

Since in each case the temperature of the one junction is kept constant at 0° C., and that when the temperatures of the two junctions is the same the thermo-electric E.M.F. is zero, all the curves must pass through the origin of co-ordinates. The temperatures at which the curves have a horizontal tangent, that is, when the E.M.F. is a maximum in one direction or the other, is the neutral temperature for the given metal taken with reference to lead. Thus the temperature of the neutral point for a platinum-lead couple is $-150°$, and that for a zinc-lead couple is $-200°$.

It is found that the curves showing the relation between the thermo-electric E.M.F., E, and the temperature, t, of the hot junction, the other junction being at 0°, is approximately a parabola. Hence, since the equation of a parabola can be written in the form

$$y = ax + \frac{b}{2}x^2,$$

where a and b are constants, the relation between the thermo-electric E.M.F. and the temperature can be expressed by a formula of the form

$$E = at + \frac{b}{2}t^2,$$

where the values of a and b depend on the nature of the metal. Those who are acquainted with analytical geometry will see that the maximum value of E occurs when t is equal to $-a/b$. Hence there exists the following relations between the constants a and b, the neutral temperature t, and the E.M.F. E' of the junction, when one junction is at 0° and the other is at the neutral temperature—

$$t' = -a/b,$$
$$E' = -a^2/2b.$$

In the following table the values of the coefficients a and b for a few metals are given. The sign of a is such that when at the hot junction the current passes from lead to the given metal a is positive, or, in other words, when a is positive the current flows in the metal considered from the hot junction to the cold. The values of the constants are so chosen

that if t is measured in degrees Centigrade the thermo-electro-motive force is given in microvolts, that is, in 10^{-6} volts.

	a	b
Copper	2.86	0.0020
Zinc	2.73	0.0035
Cadmium	3.11	0.0083
Iron	13.20	−0.0071
Nickel	−19.16	0.0072
Cobalt	−15.51	0.0190
Mercury	− 3.21	0.0042
Platinum	− 3.10	0.0051

Suppose that the thermo-electric E.M.F. between a given metal, A, and lead when the cold junction is at 0° and the hot junction is at a temperature t_1 is E, then we have seen that

$$E=at+\frac{b}{2}t.$$

Suppose now that the temperature of the hot junction is raised through a small interval δt, so that it becomes $t+\delta t$, then, if the new value of E is called $E+\delta E$, we have

$$E+\delta E=a(t+\delta t)+\frac{b}{2}(t+\delta t)^2,$$

or, since by supposition δt is very small, we may neglect the term which involves the product of the square of this very small quantity into b, which, as will be seen from the table, is itself small. Hence

$$E+\delta E=at+\frac{b}{2}t+a\delta t+bt\delta t.$$

If now we subtract the value of the E.M.F. at t from that at $t+dt$, we get

$$\delta E=a\delta t+bt\delta t.$$

That is, an increase, δt, in the temperature of the hot junction produces an increase of δE or $adt+bt\delta t$ in the E.M.F. Now the ratio of the increase in the E.M.F. produced by a small rise in the temperature of the hot junction to this increase in temperature, or, in other words, the rate of increase of E.M.F. with temperature, is called the *thermo-electric power* of the metal A with respect to lead at the temperature t. If Q is the thermo-electric power, then

$$Q=a+bt.$$

This expression shows that the thermo-electric power varies as the first power of the temperature, so that if a curve is drawn such that the abscissæ are temperatures and the ordinates are the corresponding values of the thermo-electric power, this curve will be a straight line. This is at once evident if the constant term a is subtracted, which is equivalent

to decreasing all the ordinates by the same amount, when the new ordinates will be b times the corresponding abscissæ, that is, the ordinates are directly proportional to the abscissæ, and hence the curve must be a straight line.

In Fig. 477 the lines showing the connection between the thermo-electric power and the temperature are given. Such a series of curves are called a thermo-electric diagram, and from them we can deduce the different thermo-electric properties of various combinations of metals.

Before considering this diagram in detail, we must consider two laws which have been found by experiment to hold in all thermo-electric circuits. The first of these is that if E_1 is the E.M.F. acting round a circuit composed of two metals when the temperature of the cold junction is t_1 and that of the hot junction is t_2, and E_2 is the E.M.F. when the temperature of the cold junction is t_2 and that of the hot junction is t_3,

Fig. 477.

then the E.M.F. when the temperature of the cold junction is t_1 and that of the hot junction is t_3 will be $E_1 + E_2$. But we have seen that if the temperature of the hot junction is increased by δt the E.M.F. is increased by $Q\delta t$, where Q is the thermo-electric power at the temperature t. Hence the E.M.F., when the temperature of the hot junction is t, will be the sum of the quantities obtained by multiplying the values of Q for each interval δt, starting at the temperature of the cold junction, by the interval and continuing the process up to the temperature t.

The second law is that if we have a circuit containing three metals, A, B, and C, and keep the junctions BC and CA both at the same temperature, t_1, while the junction AB is kept at the temperature t_2, then the E.M.F. acting in the circuit will be the same as that which would exist in a circuit composed of the metals A and B alone, in which one junction was kept at the temperature t_1 and the other at t_2. Thus the inter-

position of one or more intermediate metals, so long as the junctions at each end of each of these additional metals are at the same temperature, has no influence on the thermo-electric E.M.F. developed by the other two metals when their junction is at a given temperature above or below that of all the other junctions.

Returning to the thermo-diagram, since it is drawn for the different metals with respect to lead, the axis of temperatures will represent the thermo-electric line for this metal. The point where the line for any metal cuts the line for any other metal corresponds to the neutral point for these two metals; thus from the diagram the temperature of the neutral point can immediately be read off.

In order to find from the diagram the thermo-electric E.M.F. developed in a circuit, say of copper and iron, when the cold junction is at a temperature of 20° and the hot junction is at 100°, we draw the ordinates corresponding to the temperatures 20° and 100°. Then the thermo-electric power of the junction at 20° is given by the difference of the ordinates of the iron and copper lines, that is, by the length ab (Fig. 478). If the temperature of the hot junction were $20 + \delta t$, the E.M.F. acting would be $Q\delta t$, where Q is the thermo-electric power of the combination at a temperature t. But $Q\delta t$ is the area of the small strip enclosed between the iron and copper lines and the ordinates for the temperatures t and $t+\delta t$, that is, the small strip shaded in the figure. Next, if the temperature of the cold junction were $t+\delta t$, and that of the hot $t+2\delta t$, the E.M.F., in the same way, would be represented by the area of the strip between the iron and copper lines and the ordinates corresponding to the temperatures $t+\delta t$ and $t+2\delta t$, and so on. But by the law given above the E.M.F., when the temperatures of the junctions are the same as the extreme temperatures considered in these small steps, will be the same as the sum of the E.M.F.'s in the steps. Hence the E.M.F., when the junctions are at the temperatures t_1 and t_2, will be represented by the sum of all the small strips similar to ab, that is, will be represented by the area $abcd$. Since this figure is a trapezium, its area is given by $MN(\overline{ab}/2+\overline{cd}/2)$; that is, the E.M.F. acting will be equal to

Fig. 478.

$$(t_1 - t_2)(Q_1 + Q_2)/2,$$

where Q_1 and Q_2 are the thermo-electric powers of the two metals at

the temperatures t_1 and t_2. In the example taken, we get in this way from the curve $E=(100-20)(9.2+3.9)/2=524$ microvolts. If the lines of the metals intersect between the ordinates corresponding to the two temperatures, that is, if the neutral point is included between these temperatures, the areas on the two sides of the intersection must be subtracted one from the other to give the E.M.F.

500*. The Peltier Effect—The Thomson Effect.—In a thermo-electric circuit of which the junctions are at different temperatures, there is a current flowing; and we have seen in § 493 that the passage of a current through a conductor involves the expenditure of some energy, which appears as heat, according to Joule's law. The question now arises in what manner the energy necessary for the maintenance of the current in the thermo-electric circuit is supplied. This question is answered by a discovery made in 1834 by Peltier, who found that when an electric current is passed through a thermo-electric junction, *i.e.* a junction of two different metals, there will be either a development of heat at the junction or an absorption, according to the direction in which the current is passed.

The Peltier effect differs from the Joule heating already considered, in that while the Joule heating is proportional to the square of the current, and is independent of the direction of the current, the heat developed at a junction of two metals is proportional to the first power of the current, and depends on the direction of the current.

The Peltier effect can be shown, and its magnitude measured by means of the apparatus shown in Fig. 479. It consists of two glass bulbs, M and N, which are connected together by a narrow tube, in which a small drop of liquid is placed to serve as an index. A, A, and B are rods of two metals which pass through corks which fit air-tight in the tubulures of the bulbs, in such a way that the junctions m and n are at the centres of the bulbs. The portions of the rods A and B within the two bulbs are of the same size, resistance, &c., and the bulbs are also of the same size; thus the development of heat according to Joule's law is the same in both, so that, as far as this source of heat is concerned, the air in the two bulbs becomes equally warmed, and hence the pressure is the same in both of them, and the index does not tend to move in either direction. The Peltier effect will, however, be different in the two bulbs, for while in one the current will be passing at the junction from the metal A to B,

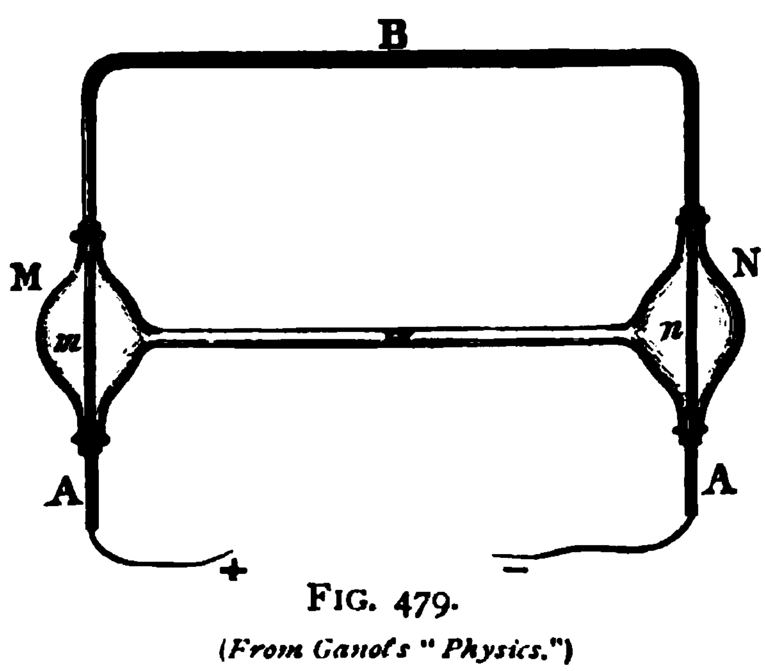


FIG. 479.
(*From Ganot's "Physics."*)

at the other junction it will be passing from B to A. Hence while heat will be developed at one junction, and the air in the bulb will be heated and its pressure increased, at the other junction heat will be absorbed, and therefore the air will on this account be cooled and its pressure decreased, causing the liquid index to move towards the junction where the heat is absorbed. If the metal A is copper and B is iron, the current being passed in the direction from copper to iron at the junction m, then heat will be absorbed at m and produced at n. On reversing the direction of the current, the direction in which the index moves will also be reversed.

In the following table the magnitude of the Peltier effect for some metals is given. The current is supposed to pass from copper to the metal mentioned in the first column, and in the second column is given the quantity of heat liberated by one ampere in one second at the junction, expressed in calories.

PELTIER EFFECT.

Metal.	Calories per Coulomb.
Iron	$-\ 1.7 \times 10^{-4}$
Platinum . . .	$+\ 0.9 \times 10^{-4}$
Silver	$-\ 1.1 \times 10^{-4}$
Zinc	$-\ 1.6 \times 10^{-4}$
Nickel	$+12.1 \times 10^{-4}$

Now let us consider a thermo-electric circuit composed of iron and copper, and suppose that one of the junctions is immersed in a mixture of ice and water, while the other junction is placed in a beaker of water at a temperature t. Then we know that a current will flow, the direction of which at the hot junction will be from the copper to the iron, while at the cold junction it will be flowing from the iron to the copper. Now when a current flows from copper to iron there is, according to Peltier's observation, an absorption of heat, while when a current flows from iron to copper there is a liberation of heat. Hence in our example there will be an absorption of heat at the hot junction, which will be supplied from the heat of the hot water, while there will be a liberation of heat at the cold junction, which will melt some of the ice. We thus see that the production of a thermo-electric current is accompanied by a transfer of heat from the source which is used to maintain the temperature of the hot junction to the refrigerator used to cool the cold junction. If then the heat given up to the refrigerator, or the condenser, as we may call it from the analogy with the thermal engines considered in the sections on thermo-dynamics, is less than the quantity absorbed from the source by

the amount of heat developed in the circuit according to Joule's law, the maintenance of the current is at once accounted for. We are thus led to look upon a thermo-electric circuit as an ordinary heat engine in which a certain quantity of heat is taken in at a given temperature; some of this heat used up by the engine in this case is first converted into electric energy, and then reconverted into heat, according to Joule's law, but, as we shall see later, might in part at least be converted into mechanical work, while the remainder of the heat taken from the source is given out to a condenser which must be at a lower temperature than the source.

In addition, a certain amount of heat will pass from the source to the refrigerator by conduction along the wires which form the circuit. Now of these three thermal processes two, namely, the heat developed, according to Joule's law, and the heat conducted along the wire, are irreversible (§ 264), for if we reverse the direction of the current in the circuit, heat will still be *developed* in the circuit, and heat will still pass from the hot end of the circuit to the cold by conduction. The heat transfer, owing to conduction, is quite independent of the quantity of electricity which passes round the circuit, and hence we shall neglect this effect when treating the circuit as a heat engine. We may, of course, if we like, render the conduction of heat quite inappreciable, by making the conducting wires forming the circuit very long. As to the heat developed according to Joule's law, this not being directly proportional to the quantity of electricity passing round the circuit, and further, since by making the resistance of the circuit sufficiently small, we may make the quantity of heat produced as small as we please, we shall also neglect this effect. Thus when we are considering the connection between the E.M.F.'s which act at the junctions, that is, the work done when unit quantity of electricity passes round the circuit, and the thermal effects which accompany the passage of the electricity, we shall only consider the reversible processes, and shall treat the circuit as a reversible engine (§ 261).

Let P_1 be the heat liberated at the cold junction of a thermo-electric circuit, of which the temperature on the absolute scale (§ 261) is T_1, when unit quantity of electricity crosses the junction, and P_2 be the heat absorbed at the hot junction which is at an absolute temperature T_2. The quantities of heat P_1 and P_2 being expressed in mechanical units, P_1 will represent the mechanical equivalent of the heat liberated at the cold junction, and P_2 that of the heat absorbed at the hot junction when unit quantity of electricity flows round the circuit. Now if W is the work done when the unit quantity of electricity flows round the circuit, we have by the equations found in § 261—

$$\frac{W}{T_2-T_1}=\frac{P_1}{T_1}=\frac{P_2}{T_2}.$$

But the work done when unit quantity of electricity flows round the circuit is equal to E. Hence

$$\frac{E}{T_2-T_1}=\frac{P_1}{T_1}=\frac{P_2}{T_2},$$

or

$$E=(T_2-T_1)P_1/T_1.$$

Hence if the temperature T_1 of the cold junction is kept constant, this reasoning would lead us to expect that the E.M.F. in the circuit would be proportional to the difference between the temperatures of the hot and cold junctions. We have, however, seen that if the temperature of the hot junction is raised, the E.M.F. will increase at first, but will eventually reach a maximum value, and when the temperature of the hot junction is as much above the neutral temperature as that of the cold one is below, the E.M.F. will be zero. From considerations of this nature Lord Kelvin was led to infer that in a thermo-electric circuit there must be other reversible thermal effects due to the current besides the Peltier effect, and these effects must account for the discrepancy between the expression we have found above on the supposition that the Peltier effects were the only reversible ones in the circuit and the observed facts. He then made a series of experiments with a view to discover such effects, and found that when a current flows along a wire of which the temperature varies from point to point, heat is liberated at a given point of the wire when the current flows in one direction, and is absorbed at this point when the direction of the current is reversed. Of course the liberation of heat here spoken of is additional to the heat liberated according to Joule's law. The relative directions of the current and of the temperature gradient, in order that there may be absorption of heat, depend on the metal. Thus in the case of copper heat is absorbed when the current is flowing from the cold part of the wire to the hot, while in the case of iron, on the other hand, heat is absorbed when the current is flowing from the hot part of the wire to the cold; that is, in the same direction as the flow of heat due to conduction in the wire. In each case reversing the direction of the current changes the absorption of heat into a liberation. This reversible thermal effect, produced when a current flows along an unequally heated conductor, is called the Thomson effect.

PART V.—MAGNETIC INDUCTION

CHAPTER XI

MAGNETIC INDUCTION

501. Intensity of Magnetisation.—We have already referred, in § 417, to the fact that a piece of iron when placed in a magnetic field becomes magnetised by induction, and we now have to investigate this phenomenon of induced magnetism in more detail.

We have defined the magnetic force at a given point in air as the force in dynes which would act on a unit pole placed at the point, and we have seen how the direction of the force which would act on a north pole when placed anywhere in the air surrounding a magnet may be mapped out by means of magnetic lines of force. We also found that in the air surrounding the magnet these lines of force ran from the north pole of the magnet to the south pole. In the case of electro-static lines of force, since there is no force exerted within a closed conductor, we did not have to consider the forms of the lines of force within a conducting body, so that a line of force *originated* at a positively electrified conductor and *ended* at the surface of a negatively electrified body. In the case of the magnetic lines of force due to an electric current (§ 472) we, however, found that they consisted of closed curves, that is, each line of force is continuous, and has neither beginning nor end. We are thus led to consider whether the lines of magnetic force due to magnets are like lines of electro-static force, originating at a north pole and ending up at a south pole, or whether they are like the magnetic lines of force of a current and are continuous curves. Suppose that a long thin steel rod is magnetised uniformly, then there will be a pole at each end and magnetic lines of force will run, in the surrounding air, from the north pole to the south pole. Now suppose that the magnet is bent into the form of a circle, the two poles being brought into contact with one another. Before the poles were brought into contact there were a number of lines of force passing through the air, but when the poles are in contact practically no lines of force pass through the air. The magnet, however, is still magnetised, for if the poles are separated, or if it is cut at any other place and the ends are separated, lines of force will at once pass through the air. We are thus led to conclude that the magnetic

lines of force exist even when they do not pass through the air, and that when the poles are in contact the lines run round the steel ring thus formed. When the poles are separated the lines of force still form continuous lines, the direction in the air being from the north pole to the south, while for the remainder of their course they run in the substance of the steel, the direction here being from the south pole to the north.

Just as in the case of electro-static lines of force we were able by means of the consideration of tubes of force to represent the strength of the electro-static field at any point, so by the consideration of tubes of magnetic force we can represent the strength of the magnetic field at every point. A unit magnetic tube of force is bounded by lines of magnetic force, and is such that the product of the magnetic force at any point of the tube into the cross-section of the tube at that point is equal to unity.

Attention must be paid to the difference between the methods in which the electro-static and the magnetic unit tubes are defined. Since the electro-static tube starts and ends at given points, we are able to define the unit tube by the quantity of electrification at the ends. Thus defined, we have seen that the product of the electric force at a point in a tube into the cross-section of the tube at that point is equal to 4π. In the magnetic case, on the other hand, the tubes being endless, we have to adopt another method of defining the unit tube, namely, that the product of the magnetic force into the cross-section should be constant. The constant usually adopted being, however, not 4π, but unity. Greater uniformity, no doubt, would be secured by altering the definition of the electro-static unit tube so as to conform to the definition of the magnetic tube in the manner given at the end of § 455. Usage having decreed that the definitions we have adopted should be used, it seems hopeless to attempt to make any change. The difference in the definitions will account for the frequent occurrence of the quantity 4π in formulæ dealing with one kind of tube, but not in the corresponding formula dealing with the other kind of tube.

If a long thin magnet is taken of which the pole strength is m and a sphere of radius r, where r is small, be described with the pole as centre, the force at any point on this sphere is m/r^2, since the other pole is at a great distance, so that it exerts no appreciable force. If a is the cross-section of a tube of force at the surface of the sphere, then by the definition of the unit tube we have

$$a . m/r^2 = 1,$$

or

$$a = r^2/m.$$

Hence, since the area of the surface of a sphere of radius r is $4\pi r^2$, the number of tubes of force which cross the surface of the sphere is $\frac{4\pi r^2}{a}$, or $4\pi m$; that is, there are $4\pi m$ tubes of force which leave the north pole of

a magnet, of which the pole strength is m. Since all these tubes of force return to the north pole through the substance of the iron, there will be $4\pi m$ tubes of force passing through the magnet, these tubes of force being all due to the magnetism of the magnet itself.

If we accept the hypothesis of molecular magnets (§ 420), the strength of the pole m either measures the strength of the molecular magnets, or, if we assume that they are of constant strength, the proportion of them which are turned with their axes in one direction. Hence for a given magnet the degree to which the steel is magnetised will depend on the number of tubes of force which pass through the substance of the magnet, that is, on the closeness of the packing of the tubes. This closeness depends on the strength of the poles of the magnet and also on the cross-section of the magnet. If then s is the cross-section of the magnet, and m is the strength of either of the poles, the degree to which the steel is magnetised is measured by $4\pi m/s$. Now if l is the length of the magnet, and its magnetic moment is M, we have $m = M/l$. Hence the degree of magnetisation of the steel is measured by $4\pi M/ls$, or since, if V is the volume of the magnet, $V = ls$ by $4\pi \frac{M}{V}$. Thus the degree of magnetisation of the steel is proportional to the quotient of the magnetic moment by the volume, and this quantity is called the ***intensity of magnetisation*** of the steel. That the value of this quantity does not depend on the form of the magnet can be seen by the following considerations. If a magnet were taken of length $l/2$ and cross-section $2s$, so that the volume was the same as before, and were magnetised so that its moment was M, the strength of each pole would be $M/2l$, that is, $2m$, and the number of tubes of force passing through the magnet would be $4\pi m \times 2$, and the number of tubes of force passing through unit area of the cross-section would be $8\pi m/2s$ or $4\pi m/s$, which is the same number as obtained before. Since the moment and the volume are the same as before, the intensity of magnetisation is the same as before, and we have just seen that the closeness of the tubes of force is the same.

The number of tubes of force which pass through the magnet is $4\pi m$, and the number of tubes per unit area of cross-section is $4\pi m/s$, but the number of tubes per unit area of the cross-section is also equal to $4\pi I$. Hence $I = m/s$. The quantity m/s, which is the pole strength per unit area, is called the surface density of the magnetisation, and from the above relation it is seen that this quantity is numerically equal to the intensity of magnetisation, I. Thus if a disc of iron, of which the area of the face is S, is magnetised transversely, the intensity of magnetisation being I, then the pole strength of either side will be IS.

502. Magnetic Induction.—We have seen that when a piece of iron is placed in a magnetic field it becomes magnetised owing to induction, and we have now to consider how the intensity of the induced magnetism depends on the conditions under which the induction takes place.

In any portion of a magnetic field which is filled with non-magnetic medium, all the tubes of force are due to causes (currents, magnets, &c.) which are external to the portion of the medium considered. If we assume, and we shall see later to what extent this assumption is justified, that in a magnetic medium the molecules of the medium are already magnetised, and that the act of magnetising any given portion of the medium consists of turning these molecular magnets in a given direction, then when such a medium is *unmagnetised* the molecular magnets will be turned in all directions. Each molecular magnet will have its tubes of force, just as a large magnet, but in the unmagnetised state these tubes of force will be turned in all directions, so that on the whole there will be as many tubes passing through any element of area, taken in the medium, in one direction as in the opposite direction. Where, however, a magnetic medium is placed in a magnetic field a certain proportion of the molecular magnets will be turned in the direction of the lines of force of the field, so that their tubes of force all point in the same direction. This effect is shown in Fig. 407, which represents the result of sprinkling iron filings over a sheet of glass on which were placed a number of small magnets, the axes of which were arranged irregularly in all directions. In Fig. 408 the corresponding figure is shown for the same magnets, but here they have been all arranged with their axes pointing in one direction. It will be observed that now there are lines of force extending to some distance from the group of magnets.

Thus within a magnetisable medium, which is placed in a magnetic field, we have to do with two sets of tubes of force—(1) those which are due to the magnetising field, and which would exist if the magnetic medium were replaced by a non-magnetic medium; (2) those due to the magnetism of the molecules of the medium itself.

Suppose that a long unmagnetised cylindrical bar of soft iron of cross-section s is placed in a uniform magnetic field of strength H, with its length parallel to the lines of force of the field. Then if the cylinder were of an unmagnetisable material, the number of tubes of force due to the field which would cross the cross-section of the cylinder is sH. Owing, however, to the fact that the cylinder becomes magnetised by induction, there will be in addition a certain number of tubes of force both within the cylinder and in the air outside, due to this induced magnetism.

Owing to the induction, poles will be induced at the ends of the iron and these poles will in general produce a force within the material of the iron which will tend to diminish the strength of the inducing field. This disturbing action of the induced poles causes a considerable complication in the consideration of the problem, and so we shall at first consider the case of a very long cylinder of comparatively small cross-section. In this case, if we confine ourselves to a consideration of the state of the iron near the middle of the cylinder, the influence of the

poles, which are by supposition at a considerable distance, may be neglected.

If m is the strength of the poles induced in the iron, then, as we have seen in the last section, there will be $4\pi m$ lines of force due to this induced magnetism. Since that end of the cylinder which points in the direction in which the lines of force of the field run becomes a north pole, the induced lines of force will run in the air in the opposite direction to the lines of force of the field, but within the iron they will run in the same direction as the lines of force of the field. The number of tubes which pass through the iron is therefore sH tubes, due to the inducing field, and $4\pi m$ tubes, due to the induced magnetisation, or $sH+4\pi m$ in all. Hence the number of tubes of force which cross unit area of the cross-section of the iron is $H+4\pi m/s$. But the intensity of magnetisation of the iron is equal to m/s, for the lines of force due to the field alone have nothing to do with the magnetism of the material, and in fact remain the same whatever the nature of the material of which the cylinder is composed. Hence if I is the intensity of the magnetism induced in the iron, the number of lines of force which cross *unit area* of the cross-section of the cylinder is $H+4\pi I$. This quantity is called the *induction*, B, in the iron, so that

$$B=H+4\pi I.$$

We have in the above spoken of the tubes of force due to the magnetising field, and to the induced magnetism which is induced in the iron, and we have defined the induction as the number of tubes of force which cross unit area at right angles to the tubes. Now although in the example we have taken we have for simplicity supposed that the cylinder of iron was placed with its length parallel to the lines of force of the magnetising field, and was entirely magnetised by induction, so that inside the iron the lines of force due to the field and those due to the induced magnetism were parallel, this is not always so. Thus if the cylinder had been placed with its length inclined to the lines of force of the magnetising field, and been permanently magnetised, the lines of force within the iron due to the permanent magnetism of the iron would not be parallel to those due to the field.

Since in the iron we have always to do with the resultant of the two sets of tubes, that is, with the induction, it is usual to speak of the tubes of induction within any magnetisable material. Thus a line of induction is a curve drawn so that it everywhere indicates the direction of the induction, that is, of the resultant of the field causing the induced magnetisation and that of the magnetisation, both permanent and induced, of the material. A tube of induction is a portion of space bounded by lines of induction, and such that the induction across every cross-section of the tube is equal to unity.

If there are no magnetisable materials in the field, then there will be

no lines of force due to induced or permanent magnetism, and the lines and tubes of induction will be the same as the lines and tubes of force. In cases where we wish to distinguish between the two sets of tubes within a magnetisable material, we shall take the terms *lines* and *tubes of force* as referring to the lines or tubes which would occupy the space if the magnetisable material were removed and were replaced by a non-magnetisable substance such as air, reserving the terms *lines* and *tubes of induction* for the resultant of the magnetising and induced fields.

In the case of a magnet placed in a region where there are no external magnetic forces, the tubes of induction within the magnet will spread out into the air from the north pole, curving round and entering the magnet at the south pole, then travelling through the substance of the magnet to the north pole.

When we come to the question as to how the magnetic induction within a mass of iron is to be measured, we are met with the difficulty that we are not able to bring the unit pole to a point within the substance of the iron without making a cavity in the iron, and the question arises, what will be the effect of such a cavity on the induction? Suppose that we return to the case of the cylinder of cross-section s, before considered, which is placed in a field of strength H, with its length parallel to the lines of force of the field, and that we cut the cylinder in two by a plane at right angles to its axis, and separate the halves by a very small amount. Then, if the gap between the parts is sufficiently small, practically the whole of the tubes of induction which leave one portion will enter the opposing face of the other portion. Where the tubes of induction leave the face at one side of the air gap a north pole will be produced, while on the other side of the gap, where they enter the iron again, there will be a south pole. These poles will, however, be so very near that, being of the same strength, they will exactly neutralise each other's effect, except in the air gap. Let F be the value of the field in the gap, that is, the force which would act on the unit pole if placed there, then, as the field will be uniform, the number of tubes of force which cross the gap will be sF, for the cross-section of the gap is s. Now the number of tubes of induction which enter the air gap from the iron on the one side and re-enter the iron on the other side is sB, where B is the induction within the iron. Hence, since each of the tubes of force in the air gap is the continuation of one of the tubes of induction in the iron, we must have the same number of each, that is, $sB=sF$ or $B=F$. The magnetic induction is therefore equal to the strength of the magnetic field in a small crevasse cut in the iron at right angles to the lines of induction in the iron.

503. Magnetising Force.—When considering the case of the cylinder of iron magnetised inductively in the last section, we assumed for simplicity that the length was very great, and were thus able to neglect the

magnetic force within the iron due to the poles which are induced at the ends, and we have now to consider what influence these will have on the results obtained.

If the length of the bar is finite, so that the influence of the poles induced at the ends has to be taken into account, the magnetising force at any given point of the bar will consist of two parts, that due to the field when the iron is not present, and that due to the poles which are induced at the ends. If it were possible to keep the induced poles in the place which they occupy and yet remove the iron, then the field due to the poles would be opposite in direction to that due to the inducing field, for the induced north pole is turned towards the direction in which the lines of force of the field run and the lines of force due to the two induced poles run from the north pole to the south. The resultant of the two fields is the field which is actually efficacious in producing the induced magnetisation of the iron, and is called the magnetising force. Since the operation of removing the iron without moving the poles is impossible, we have to consider some other method of measuring the magnetising force.

Suppose a long thin cylindrical cavity to be made in the iron with its axis parallel to the direction in which the iron is magnetised. Then, since the axis of the cylinder is parallel to the direction of magnetisation, all the molecular magnets will be arranged parallel to the sides of the cavity, so that there will be no free poles developed on the sides but only on the ends. Hence, if the length is very great compared to the diameter, the force due to these free poles, produced by the magnetisation of the iron on the ends of the cavity, will be negligable, and a unit pole placed inside the cylindrical cavity will experience a force which is due solely to the inducing field and to the poles induced at the ends of the iron cylinder. In other words, the force acting on the unit pole will be equal to the magnetising force which is acting inductively on the iron and which produces its induced magnetisation, and not in any way due to the influence of the neighbouring iron. Hence in the equation $B = H + 4\pi I$, in the case where the iron is of such a length that the poles induced at the ends exert an appreciable magnetic force at the point in the iron considered, H must be taken to be the force which would act on the unit pole if placed inside a very long cylindrical cavity at the given point, the axis of the cavity being parallel to the direction in which the iron is magnetised.

We have seen that the effects of the poles induced at the ends of the iron is to reduce the magnitude of the magnetising force, so that if the external field is removed the force due to these induced poles will tend to demagnetise the bar. The longer the bar, the smaller will be their demagnetising force, so that it is better, when making permanent magnets, to use long thin bars of steel than short and broad ones. In most experiments on the strength of the magnetism induced in iron rods the

length is taken so great that the demagnetising force, due to the poles induced at the ends, is negligable, so that the magnetising force will be equal to the strength of the field before the introduction of the iron, and the argument in the last section will hold good.

504. Susceptibility—Permeability.—If the magnetism of a piece of iron is entirely induced, the magnetisation and the induction are parallel to the direction of the magnetising force; and if H is the magnetising force, B the induction, and I the intensity of magnetisation, we have

$$B = H + 4\pi I.$$

The ratio of the intensity of the induced magnetisation to the magnetising force, or I/H, is called the *magnetic susceptibility*, and is generally indicated by the letter k.

The ratio of the induction to the magnetising force, or B/H, is called the *magnetic permeability*, and is generally indicated by the letter μ. Hence

$$I = kH \text{ and } B = \mu H.$$

Substituting these values in the equation for B, we get

$$\mu = 1 + 4\pi k.$$

The permeability, μ, has the same relation to magnetism as the specific inductive capacity has to electro-statics. There is, however, this important difference, that while the specific inductive capacity of a dielectric is independent of the electro-static force acting and of the values of the force which have previously been acting, the contrary is the case with the permeability. Thus in the case of iron the value of the permeability, at any rate for magnetising forces above about 0.02 *c.g.s.* units, depends not only on the value of the magnetising force actually acting, but also on what magnetising forces have previously acted on the iron.

In Fig. 480 is given a curve showing the relation between the magnetising force and the magnetic induction for a particular sample of iron. It will be seen that the curve for small magnetising forces is nearly straight; after this it bends sharply upwards, and then gradually becomes flatter and flatter till, for large values of the magnetising force, it is again straight.

In the figure the scale adopted for B is very much smaller than that for H, since the changes in B are so very much greater than are those in H. If the same scale is adopted for the two, the curve for large values of H is a straight line inclined at 45° to the two axes, indicating that the change in B is equal to the change in H.

If we consider any point P on the curve, then, since the permeability is the ratio of the induction to the magnetising force, the value of the permeability for the magnetising force corresponding to P is given by

the ratio of the ordinate of P to the abscissa, or, what comes to the same thing, by the tangent of the angle θ between the line joining P to the origin and the axis of H.

Since for very small values of H the curve is straight, the tangent of the angle θ is constant, that is, the permeability is constant. As H increases θ increases, and hence the permeability also increases. At the point Q on the curve the value of θ is a maximum; thus for the corresponding value of H, that is, 4.8 *c.g.s.* units, in the case of the sample of iron for which the curve is drawn, the permeability is a maximum. For greater values of H, θ, and therefore also the permeability, decreases.

Since for high values of H the curve is a straight line inclined at 45° to the axis of H, for all points on this portion the curve we have $B=H+a$,

FIG. 480.

where a is a constant. Comparing this relation with that between B, H, and I, namely, $B=H+4\pi I$, we see that it indicates that $4\pi I$ has become constant. Hence the intensity of magnetisation does not increase indefinitely as the strength of the magnetising field is increased, but eventually reaches a constant value. Under these circumstances the iron is said to be saturated. In the case of a specimen of soft iron examined by Ewing, saturation was produced by a magnetising force of less than 2,000 *c.g.s.* units. The magnetising force required to saturate steel is very much greater than that required in the case of soft iron. Thus Ewing found that a magnetising force of 15,000 units was not sufficient to saturate one specimen of steel. In Fig. 481 the curve showing the relation between the magnetising force and the permeability is given for the sample of iron for which the curve in Fig. 480 is drawn.

It will be noticed how the permeability is at first almost constant, then increases rapidly to a maximum, then decreases—at first rapidly, then more slowly. The curve showing the connection between the intensity of

FIG. 481.

magnetisation and the magnetising force is similar to the $B-H$ curve, since the intensity of magnetisation is practically equal to $B/4\pi$, the values of H being very small compared to B.

505. Effects of Temperature on the Magnetic Properties of Magnetic Metals.—When a piece of iron or steel is heated to a bright red it loses its power of becoming magnetised, or, if permanently magnetised, all this permanent magnetism will be lost. A similar change takes place in the case of nickel and cobalt.

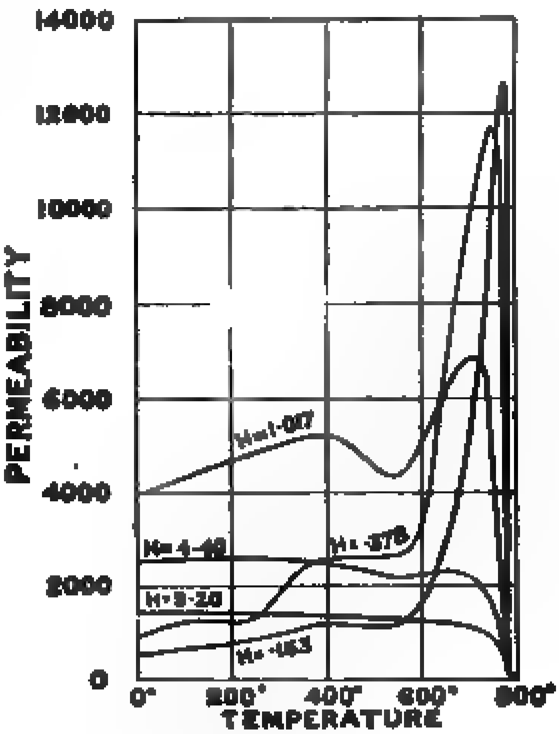


FIG. 482.

The temperature at which a particular sample of a magnetic metal loses its magnetic properties is called the critical temperature for that metal. This temperature must not be confused with the critical temperature in the case of gases, which has reference to an entirely different property (see § 232).

With small values of the magnetising force the loss of the magnetic properties of soft iron, as the temperature reaches the critical point, is much more sudden than with strong magnetising forces. In Fig. 482 the

relation between the permeability of soft iron and the temperature, as obtained by Morris, for different magnetising forces is shown.

For low magnetising forces it will be seen that the permeability increases slowly with rise of temperature up to a temperature of about 600° ; the increase then becomes very much more rapid, till at about 750° the curve becomes almost vertical. The permeability reaches a maximum value for a temperature of 775°. Above this temperature there is a sudden decrease of the permeability, and at a temperature of 785° the permeability is practically unity, that is, the iron has lost its magnetic properties. For large magnetising forces (of 4.0 *c.g.s.* units and over) there is no increase of the permeability as the temperature increases, and the decrease of the permeability as the critical temperature is approached is very much more gradual and commences at a temperature of about 670°.

Similar results are obtained in the case of steel, although the loss of magnetic properties is not so sudden as in the case of wrought iron.

For nickel the critical point is at about 300°, and the general

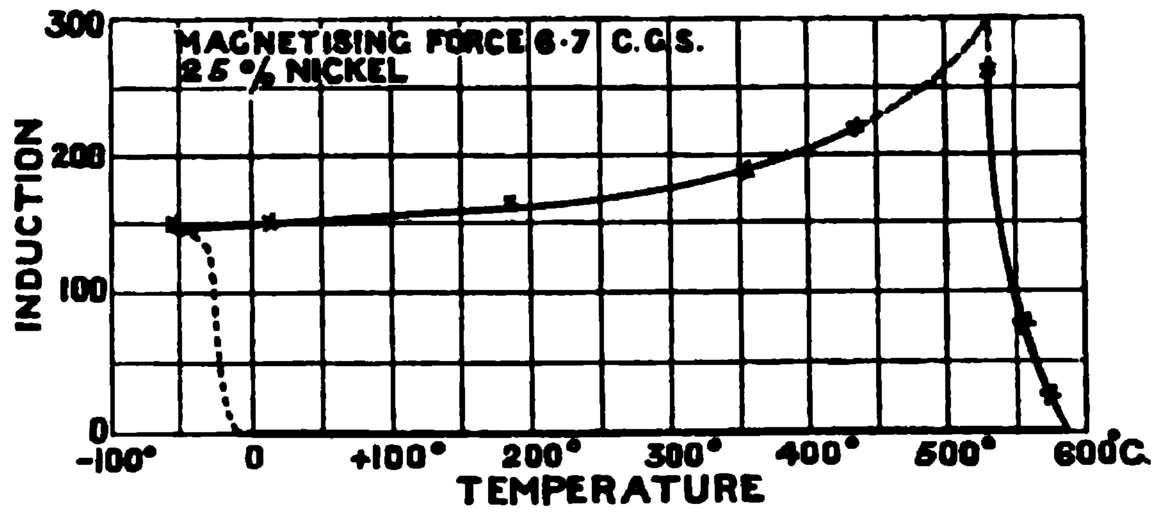


FIG. 483.

character of the changes which go on with rise of temperature are similar to those in the case of iron.

Some very remarkable results were obtained by Hopkinson for a specimen of steel containing 25 per cent. of nickel. Under ordinary circumstances this alloy is unmagnetisable at ordinary temperatures ; if, however, it is cooled down to a temperature a few degrees below zero, it becomes magnetisable. If now the temperature is raised, the magnetic property is retained till a temperature of 580° is reached, when it again becomes non-magnetic, and remains in this state till it has been again cooled below zero. The connection between the induction produced by a magnetising force of 6.7 *c.g.s.* units and the temperature is shown in Fig. 483. It will be seen that at temperatures between 0° and 580° this alloy can exist in two distinct magnetic conditions, both of which are stable, one condition changing into the other only when the temperature of the body passes through one of two fixed values.

In the case of iron the critical point is marked by the sudden change

of several other of its physical properties in addition to the change in its magnetic condition. Thus there is a rapid change in the rate at which the electrical resistance of iron changes with temperature near the critical point, as well as a change in the thermo-electric properties of the substance. That some sudden change takes place in the physical state of the substance at this temperature is shown very clearly by the phenomenon of calorescence. If a piece of iron or steel is heated to a bright redness and then allowed to cool slowly, it is found that the temperature gradually falls till the critical temperature is reached, when the rate of cooling becomes very much less, and in some cases the temperature, at any rate of the surface layers, may even rise although the loss of heat by radiation is going on all the time. This effect may be compared to the check in the rate of the fall of the temperature of a body which is changing its state, such as when a mass of metal solidifies, so that this evolution of heat at the critical point indicates that some considerable change in state occurs at this temperature. The evolution of heat in the case of hard steel is so great that it produces a visible increase in the brightness of the cooling mass of metal.

506. Hysteresis.—We have already referred to the fact that the permeability of a given sample of iron depends not only on the magnetising force to which it is at the time exposed, but also depends in some measure on the nature of the magnetising forces to which it has been previously exposed. In order to avoid the discussion of this effect in the two previous sections, we have always assumed that we started with a specimen of iron which had not been previously magnetised, and that the magnetising force was then gradually increased to the maximum value, the intermediate values of the induction, &c., being measured as the magnetising force was being increased.

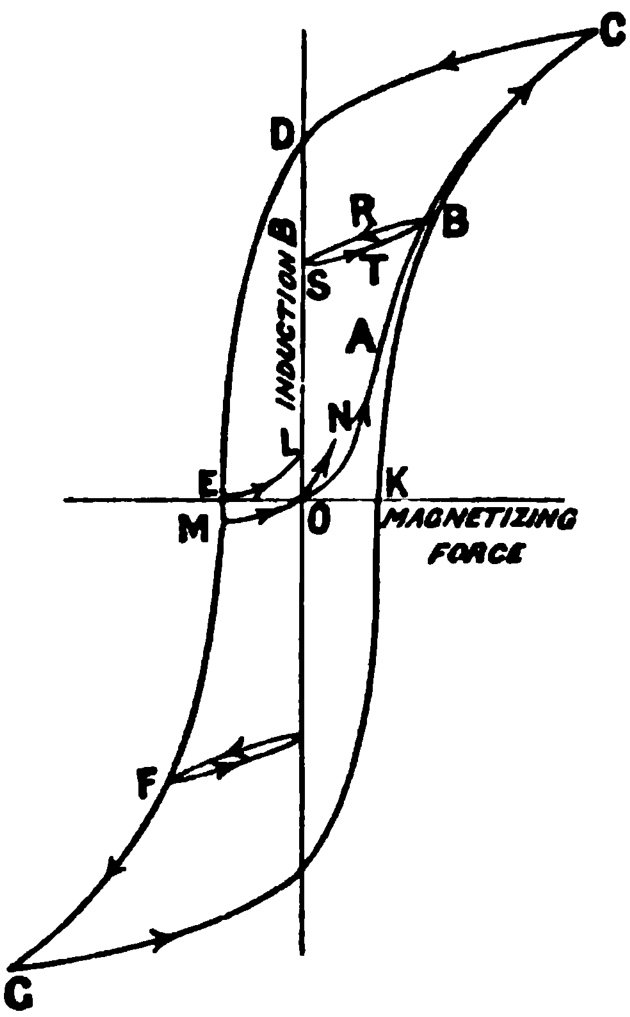


FIG. 484.

Suppose now that, starting with an unmagnetised bar of iron, we gradually increase the magnetising force and determine the corresponding values of the induction, we shall in this way obtain a curve OAC (Fig. 484) similar to the curve given in Fig. 481. If now when the point C is reached the magnetising force is gradually decreased and the value of the induction is again measured as H decreases, it will be found that the curve obtained does not coincide with the curve obtained with increasing values of H, but has the form CD.

Thus when the magnetising force is zero, the induction, instead of being zero, has a value $\overline{OD}$. This is of course due to the coercive force of the iron referred to in § 417.

If now the direction of the magnetising force is reversed, the curve DEG will be obtained; while on decreasing the magnetising force to zero, and then starting with it in its original direction, the branch GKBC of the curve will be obtained. It will be seen, by a study of this curve, that in all cases the magnetisation appears to lag behind the magnetising force, and to this phenomenon the name *hysteresis* has been applied. If, after the value of *H* has again reached its maximum positive value, it is again decreased to the same negative value, then back to the extreme positive value, the curve obtained will be very nearly, if not exactly, coincident with the curve CDEGKC. The magnetising force represented by $\overline{OE}$ or $\overline{OK}$ represents the force required to deprive the bar of its residual magnetisation. It must, however, be remarked that the condition represented by the points E and K, at which the induction is zero, is quite different from the condition at O, before any magnetising force had been applied. If when the bar is in the condition represented by E the force is reduced to zero, the induction would become positive, and the condition would be represented by the point L. Even if the magnetising field is reversed at M and then decreased to zero, so that the condition of the iron is represented by the point O, where both the force and the induction are zero, the condition of the iron is different from what it was at the start, for if the magnetising force be gradually applied, the B-H curve, as these curves showing the relation between B and H may be called, is now along ON and not along the original curve OAB.

If, after the iron has reached a condition represented by the point B, the magnetising force is gradually decreased to zero, the curve BRS is obtained, OS as before representing the residual magnetism. If now the magnetising force is gradually increased in the same direction as before, the curve STB will be obtained. Thus in this case also the B-H curve, when the value of H is taken through a cycle of values, encloses a loop. Since work has to be done to increase the induction in a piece of iron, and the greater the existing induction the greater the work that has to be done to increase the induction by a given amount, and that during a cycle a greater magnetising force has to be used to obtain a given induction while we are magnetising the iron than that which corresponds to the same induction when the magnetising field is decreasing, it follows that more work is done during the time that the rod is being taken from G to C than is done by the magnetism of the rod while it is passing along CEG. Hence a certain amount of work has to be done to carry the rod round the cycle represented by the curve, and it can be shown that this amount of work is represented by the area of the loop included by the curve. The energy expended in

doing this work appears as heat, which is developed in the iron as a consequence of the changes in its magnetisation.

507. Ewing's Molecular Theory of Magnetism.—Since, as we have seen in § 420, if we break up a magnet, each of the parts into which it is broken, however small they may be, is magnetised, we are led to look upon magnetism as some condition of the molecules, and that the phenomenon of magnetism consists of these molecular magnets being placed with their poles pointing in the same direction, the end poles being the only ones which are free to cause any external effect. Assuming that in a magnet the individual molecules of a magnet are themselves magnets, there are two hypotheses open to us; either we may suppose that in the unmagnetised state of the bar the molecules themselves are unmagnetised, and assume that when the bar is magnetised the molecules in some way or other becomes magnetised, a supposition which does not in any way assist us in explaining any of the phenomena we have been considering in the previous sections; or we may assume that the individual molecules are always magnetised, and that what happens when the bar is magnetised is that these molecular magnets are turned, so that their like poles are turned in the same direction. On this hypothesis an unmagnetised bar is one in which the molecular magnets are arranged with their axes turned in all directions, there being as many with their axes in any one direction as in any other, and that the action of magnetising the bar consists of turning a certain proportion of the molecules so that their magnetic axes all point the same way; the greater the number of molecules in the unit volume which are in this way turned into line, the greater the intensity of magnetisation of the bar. This theory of magnetism, which was first due to Weber, has been worked out by Ewing, who has given an explanation as to the causes of the forces which must be assumed to exist to account for the fact that all the molecular magnets do not set themselves with their axes in the direction of the magnetising force, however small this force may be; and, further, is capable of explaining most of the phenomena of magnetism, such as saturation, hysteresis, &c.

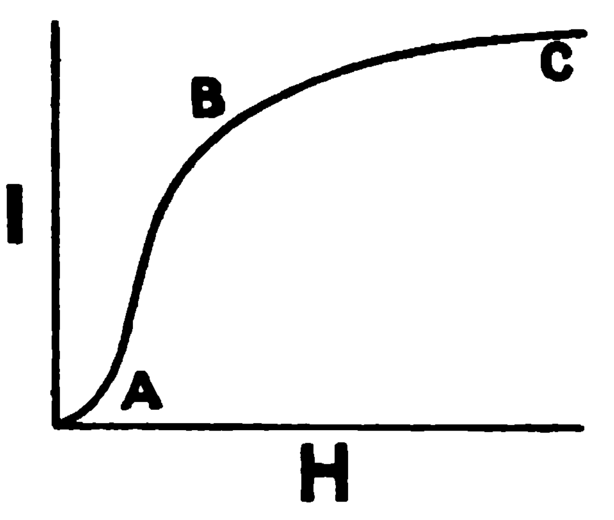


FIG. 485.

The main features which any theory of magnetism has to explain are hysteresis, saturation, and the three characteristic portions into which the curve showing the connection between the magnetising force and the intensity of the induced magnetisation can be divided. Such a curve for iron is shown in Fig. 485, and consists of three parts, A, B, and C, which mark three distinct stages in the process of magnetisation. In the first stage, A, the susceptibility I/H, is small, the curve starting off at a small inclination to the axis of H. In the

second stage, B, the susceptibility increases very rapidly, that is, a small increase of the magnetising force produces a relatively large increase in the induced magnetisation. In the third stage, C, the increase of the intensity of magnetisation with increase of H is slow, and for very large values of H practically nil. There is also a marked difference as regards hysteresis between the sections of the curve. In the first section, on the removal of the magnetising force, the iron loses nearly all its induced magnetism, there being hardly any hysteresis. In the second portion, however, on the removal of the magnetising force the iron is able to retain a considerable proportion of its magnetism, while in the third stage the amount of the residual magnetism is hardly greater than in the second stage.

As an introduction to Ewing's theory, let us consider the case of two small magnetic needles, which are supported on fixed pivots near each other, but not so near that the poles of the needles can come in contact. If there is no external field these needles will take up a position such as that shown at (*a*), Fig. 486, in which the axes of the magnets are parallel to the lines joining the pivots. Suppose now that an external magnetic force, which is at first weak, acts in the direction of the arrow H. As a result of this weak external field the magnets will be slightly deflected, and on the removal of the field they will return to their original positions. This corresponds to the stage A of the I-H curve. If, however, the value of H is increased a stage will at length be reached when the magnets will suddenly fly round into the positions shown at (*c*). On further increasing H, the magnets will set themselves more and more nearly with their axes parallel to the direction of the field. If now the value of H is gradually decreased, the inclination of the axes of the magnets to the direction of H will gradually increase, until for some value of H, which will be less than that for which the sudden swing round occurred, the magnets will suddenly return to the position shown at (*b*). We have, therefore, in this excessively simple arrangement, three distinct stages; the first, in which the magnetising field only produces a small deflection, which is such that on the removal of the magnetising force the deflection becomes zero. Secondly, a stage where the magnets reach an unstable position and then suddenly swing round into a new configuration, and where this configuration does not break up until the deflecting force reaches a value smaller than that for which the unstable condition was reached when H was increasing; and,

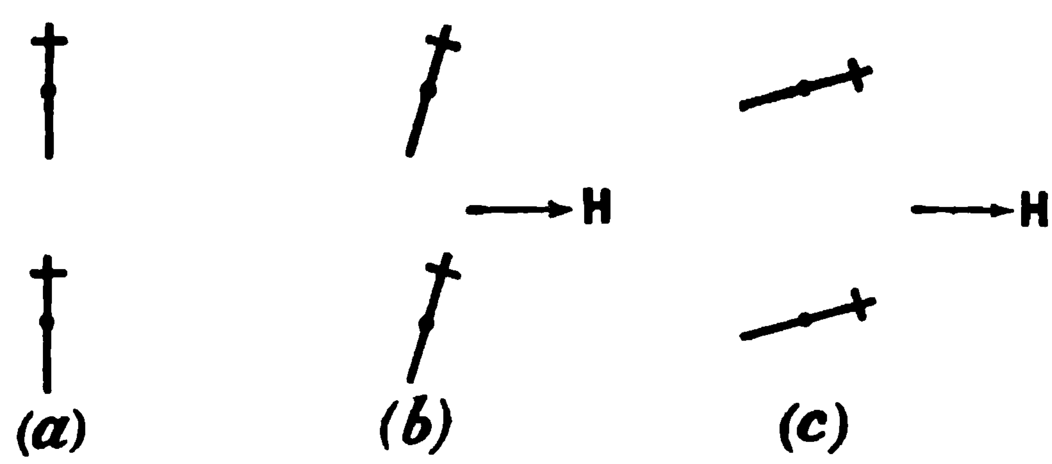


FIG. 486.

thirdly, a stage when increase of H only produces a small increase in the alignment of the magnets. Thus with only two magnets an indication of the chief peculiarities of the magnetisation curve can be obtained.

By considering much larger numbers of such pivoted magnets a much nearer approach to the phenomena actually found in the case of the magnetisation of a magnetic metal can be obtained. We have, however, said enough to indicate the line of argument by means of which Ewing supports his theory, and for further details we must refer the reader to his original papers on the subject.

In order to account for the heat developed in iron, due to hysteresis, when it is taken through a cycle of magnetisation, Ewing supposes that, on the decrease of the magnetising force, the molecular magnets return towards their undisturbed positions, and in doing so acquire kinetic energy, so that instead of immediately coming to rest they will execute oscillations about their position of rest till the kinetic energy thus acquired is converted into heat due to the currents induced in neighbouring molecules (see § 516).

In the above molecular theory of magnetism no supposition has been made as to the cause of the molecules being magnets. To account for this Ampère put forward the hypothesis that the magnetism of the molecule was really the field of an electric current which circulates continuously within it. In order to account for the fact that these molecular currents must continue without diminution, it is necessary to suppose that the molecule offers no resistance to the circulation of these intra-molecular currents such as occurs when a current passes between one molecule and another in the phenomenon of conduction.

The direction in which the Ampèrian molecular currents must be supposed to circulate can at once be obtained from either of the rules given in § 471. Since if we face a north pole the lines of force run from the pole towards us, and in a circle conveying a current the lines of force flow towards the spectator when the current circulates in the anticlockwise direction, it follows that when facing the north pole of a magnet the molecular currents must circulate in the anticlockwise direction.

In the same way, if we suppose that the earth's magnetic field is due to currents circulating round the earth, since the pole near the geographical north pole is what we call in magnetism a south pole, it follows that the currents must flow in the east to west direction, that is, in the same direction as the apparent motion of the sun.

508. Paramagnetic and Diamagnetic Bodies.—Iron, nickel, and cobalt, the so-called magnetic metals, are materials in which the permeability is greater than unity, that is, greater than the permeability of air. In addition to these bodies there are others in which the permeability is only very little greater than that of air. All these substances are classed together as paramagnetic bodies. The great majority of substances, however, have permeabilities less than that of air, and are called diamagnetic.

The extent to which bodies exhibit diamagnetism is, however, very much smaller than the paramagnetism of iron, nickel, and cobalt. Thus bismuth, the most strongly diamagnetic body known, has a permeability of 0.9998, while the permeability of iron under certain conditions is as high as 2000.

If a rod of a diamagnetic material is introduced into a magnetic field, it will become magnetised by induction, but the poles will be in the opposite direction to what they would be in the case of a paramagnetic body, so that the south pole is turned towards the direction in which the lines of force of the magnetising field are running.

The fact that in diamagnetic bodies the permeability is less than it is in air means that the induction, B, through the body is less than the value of the field which would exist if the body were removed. This can only hold if the tubes of force due to the magnetism induced in the body run, within the body, in the opposite direction to the tubes of force of the field. In order that the tubes of force due to the induced magnetism of the body may, in the body, run in the opposite direction to the tubes of force of the inducing field, a north pole, that is, a place where tubes of force leave the body, must be formed at the end of the body which is turned towards the direction from which the tubes of force of the field enter the body. Hence, when a diamagnetic body is introduced into a uniform magnetic field, the lines of force within the space occupied by the body are fewer than there would be in this space were the body removed. On the other hand, outside the body the lines of force will be more closely packed than they would be in the absence of the body, for outside the body the tubes of force due to the induced magnetisation of the diamagnetic body are, on the whole, in the same direction as the tubes of force in the field. In Fig. 487, the lines of force of a uniform field in which a sphere of a strongly diamagnetic material has been introduced are shown. The corresponding case, where the sphere is composed of a paramagnetic material, is shown in Fig. 488. It will be noticed how in this case the lines of force crowd into the sphere, and are more widely spaced in the region outside the equatorial portion of the sphere.

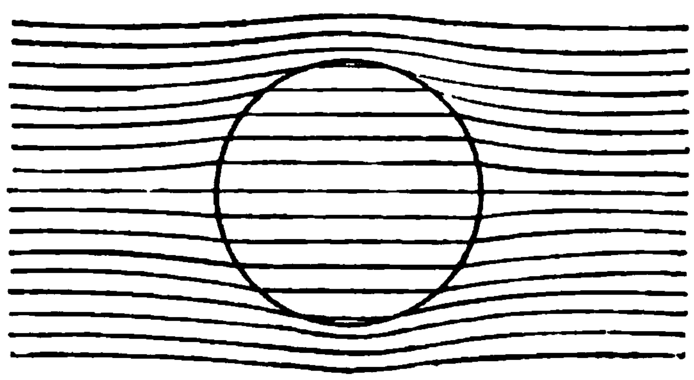


FIG. 487.

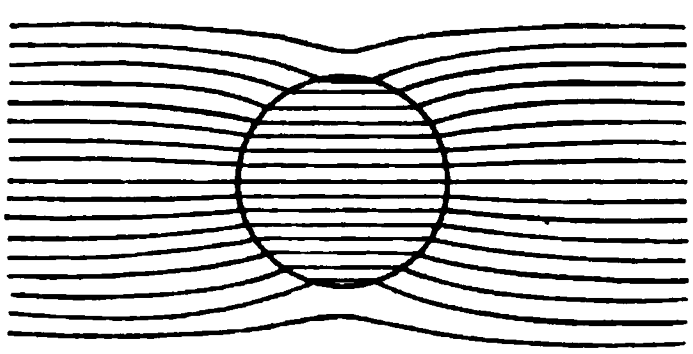


FIG. 488.

Suppose that a bar of a paramagnetic material AB (Fig. 489) is placed in a magnetic field of strength H, the direction of which makes with the length of the bar an angle θ. We

may resolve the field H into a component $H \cos \theta$ parallel to the axis of the cylinder, and a component $H \sin \theta$ perpendicular to the axis. If I_1 is the intensity of the magnetisation parallel to the axis induced by the component $H \cos \theta$, and if k is the susceptibility of the iron, and the length of the cylinder is so great that the demagnetising force due to the induced poles can be neglected, we have $I_1 = kH \cos \theta$. If the length of the cylinder is l and its cross-section is s, the volume is sl, and since the magnetic moment of a magnet is equal to the product of its volume into the intensity of magnetisation, the moment of the cylinder due to the magnetisation induced by the component of the field parallel to the axis is $sl . kH \cos \theta$. Now in § 425 it was shown that the couple acting on a magnet, of which the moment is M, tending to turn it into parallelism with a field of strength H, when its axis makes an angle θ with the direction of the field, is $MH \sin \theta$. Hence the couple, due to the magnetism induced by the component parallel to the axis, tending to turn the cylinder is $slkH^2 \cos \theta \sin \theta$.

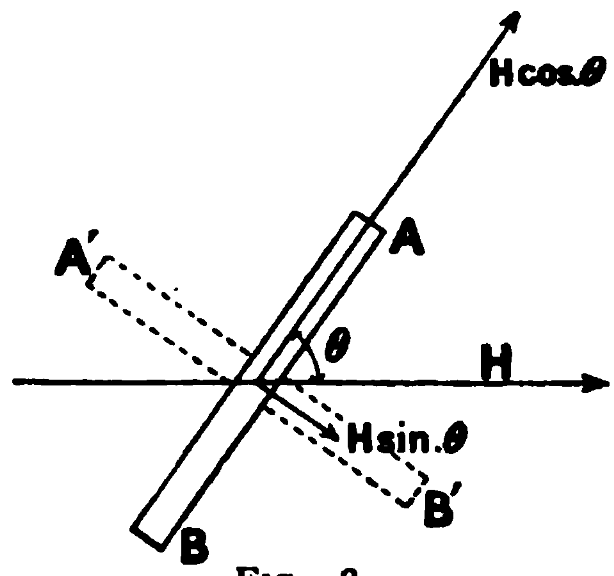


FIG. 489.

The component of the magnetising field at right angles to the axis will also induce a magnetisation in its own direction. In this case the magnetising force will be much less than $H \sin \theta$, on account of the demagnetising force exerted by the ends, and it has not yet been found possible to calculate exactly what this effect will be in such a case as that we are considering. If we consider the bar as a spheroid, of which the major axis is very much greater than the minor axis, the intensity of the transverse magnetisation can be shown to be given by

$$I_2 = \frac{k}{1 + 2\pi k} H \sin \theta.$$

Hence the magnetic moment is

$$\frac{lsk}{1 + 2\pi k} H \sin \theta,$$

and the turning moment exerted by the field is

$$\frac{lsk}{1 + 2\pi k} H^2 \sin \theta \cos \theta,$$

and is in the opposite direction to the moment due to the longitudinal magnetisation. The total turning moment is therefore given by

$$\left\{ k - \frac{k}{1 + 2\pi k} \right\} lsH^2 \sin \theta \cos \theta.$$

In the case of iron in a magnetising field about equal to that of the

earth in these latitudes k is about 30, so that the term $\frac{k}{1+2\pi k}$ is equal to 1.6, and so is small compared to the term k. The cylinder of iron thus tends to set itself parallel to the direction of the field.

In the case of a diamagnetic body the value of k is so small that $1+2\pi k$ is practically unity, and the term $\frac{k}{1+2\pi k}$ is very nearly equal to k, so that in a *uniform* field there is no measurable directive force exerted upon even a cylinder of bismuth ($k = 0.6\ 10^{-6}$). The manner in which a diamagnetic cylinder will set itself in a very strong magnetic field can, however, be at once foreseen. Since there will be a south pole induced at A (Fig. 489), and a north pole at B, the cylinder will tend to turn round in the anticlockwise direction. When it got beyond the position where its axis was at right angles to the direction of the field the polarity would be reversed, so that in such a position as A′B′ there would be a north pole at A′ and a south pole at B′, and hence the cylinder would tend to return into the position where its axis is at right angles to the direction of the field. Diamagnetic bodies therefore tend to turn, so that their longer axis is at right angles to the direction of the field.

Solids are not the only bodies which exhibit magnetic properties; thus oxygen and some solutions of iron salts are paramagnetic, while water and alcohol are diamagnetic.

By means of these liquids it can be shown that the direction in which a cylindrical tube filled with, say, a paramagnetic liquid tends to set itself depends on the susceptibility of the surrounding medium. Thus a tube containing a weak solution of ferric chloride will in air or water set itself parallel to the direction of the field, since its susceptibility is greater than that of either air or water. If, however, it is surrounded by a stronger solution of ferric chloride, it will behave like a diamagnetic body and set itself with its length perpendicular to the direction of the lines of force of the field. This effect is at once explainable if we consider that when the tube containing the weak solution is placed in the stronger solution, since the permeability of the contents of the tube is less than that of the surrounding medium, the induction through the tube will be less than that which would exist if the tube were removed, and the tube is practically diamagnetic with respect to the surrounding stronger solution.

It is therefore evident that in order to account for diamagnetism it is not necessary to assume that these bodies have a negative susceptibility, but only that their susceptibility is less than that of air, or, since the susceptibility of air and of a vacuum are very nearly the same, less than that of a vacuum. Since the susceptibility and the permeability are related by the equation

$$\mu = 1 + 4\pi k,$$

and that for the most diamagnetic body known the susceptibility is less than $1/4\pi$, the permeability will in all cases be greater than zero.

•PART VI.—ELECTRO-MAGNETISM

CHAPTER XII

FORCES ACTING ON CONDUCTORS CONVEYING CURRENTS

509. Force acting on a Straight Conductor conveying a Current when placed in a Magnetic Field.—If a straight conductor, in which a current is flowing, is placed in a magnetic field, so that it is at right angles to the lines of force of the field, then, owing to the magnetic field due to the current, the distribution of the lines of force of the field will be altered. In Fig. 490 are shown the lines of force due to a conductor which is perpendicular to the plane of the paper and passes through the point A when placed in a uniform magnetic field in which the lines of force ran parallel to the line CD. Remembering that we have every reason to suppose that there exists a tension along the lines of force, and a pressure at right angles, while the lines of force act as if they were connected with the body by which they are produced, it is evident that, as a result of the crowding of the lines of force on one side of the conductor, and their separation on the other, the conductor conveying the current will be acted upon by a force in the direction of the arrow.

If the current flows downwards, the lines of force are circles which run in the clockwise direction, and at the upper part of the diagram they strengthen the magnetic field, since they run in the same direction as the lines of force of the field. In the lower part of the diagram the lines of force due to the current and to the field are in opposite directions, and therefore the resultant magnetic field is the difference of the fields due to the two causes. The direction in which the conductor tends to move is therefore at right angles to the direction of the lines of the field, and towards the part of the field where the lines of force due to the current are in the opposite direction to the lines of force of the field. Since the direction of the lines of force of the current can at once be remembered by one of the rules given in § 472, the direction of the force acting on a conductor in a magnetic field can at once be remembered. Fleming has given a convenient rule for remembering the direction in which a conductor conveying a current in a magnetic field will tend to move. If the index finger of the *left* hand is held pointing in the direction of the lines

of force of the field, and the middle finger in the direction of the current, the conductor will tend to move in the direction of the outstretched

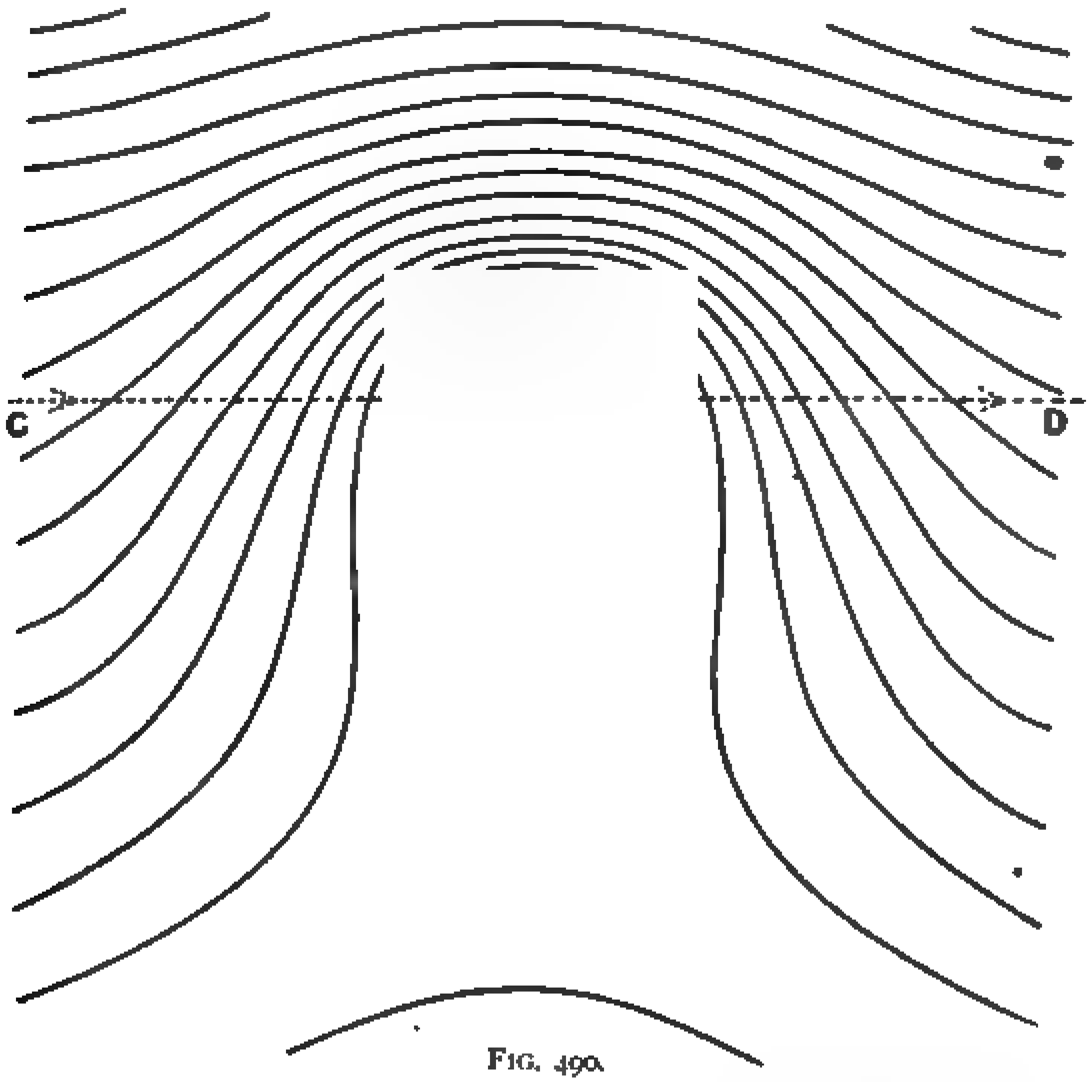


FIG. 490.

thumb, and at right angles to the lines of force of the field. A study of Fig. 491 will make the matter clear. Thus a vertical wire in which a current was flowing downwards would, on account of the earth's horizontal component, be acted upon by a force tending to move it in an easterly direction. In this case, according to our rule, the left hand must be held with the index finger pointing towards the north, since the lines of force of the horizontal component

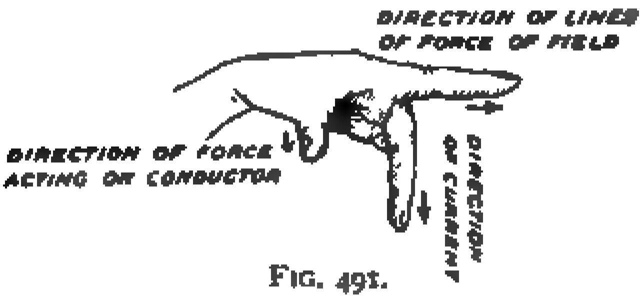


FIG. 491.

of the earth's field run from south to north, and with the middle finger pointing downwards. The outstretched thumb will then point towards the east.

Ampère, who made a lengthy series of experiments on the forces acting on conductors in which currents are flowing, showed that if a conductor of length l is traversed by a current of C *c.g.s.* units, and is placed at right angles to the lines of force of a uniform magnetic field of strength H, the force acting on the conductor will be equal to lCH. If the current is measured in amperes, then, since the ampere is one-tenth of a *c.g.s.* unit, the force will be one-tenth of the above.

If the conductor is not at right angles to the direction of the lines of force of the field, in calculating the force we must resolve the field into two components, one perpendicular to the direction of the current, and the other parallel. Then the component parallel to the direction of the current will produce no force on the conductor, and the force due to the other component is calculated by the formula given above.

510. Force acting on a Rectangular Coil conveying a Current when in a Magnetic Field.—As an example of the application of the formula given in the last section, we may calculate the force acting on a rectangular coil when placed in a uniform magnetic field. Suppose that the field is of strength H, and that the lines of force are horizontal, and run from south to north. Let the coil, ABCD (Fig. 492), consist of a single turn of wire in the form of a rectangle of length l and breadth b, and let it be placed with its plane in the vertical plane parallel to the direction of the field. Since the top and bottom of the rectangle, AB and CD, are parallel to the direction of the lines of force of the field, they will experience no force. If a current of C *c.g.s.* units is flowing round the rectangle,

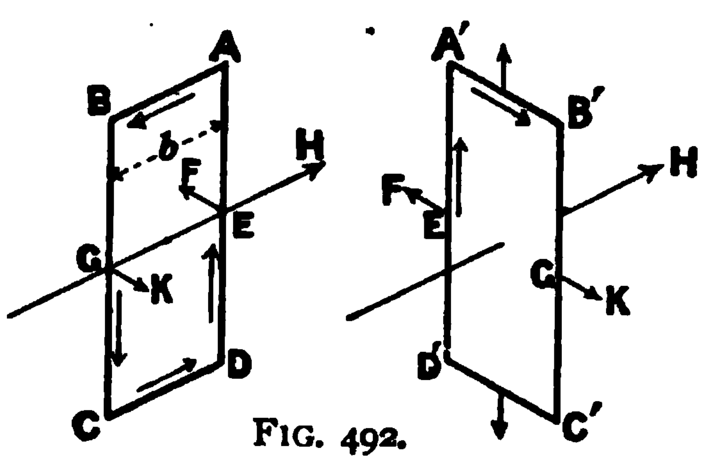


FIG. 492.

so that its direction in AB is from A to B, the vertical side AD will be acted upon by a force lCH in the direction of EF. In the same way the vertical side BC will be acted upon by an equal force in the direction GK. The resultant of these forces, since they are equal and opposite parallel forces, is a couple, of which the magnitude is $b.lCH$, tending to turn the rectangle round in the anticlockwise direction, when looked at from above.

Next suppose that the rectangle is allowed to turn round under the influence of this couple into the position A′B′C′D′, in which its plane is perpendicular to the direction of the lines of force of the field. In this position the top A′B′ is now perpendicular to the lines of force, and therefore experiences a force. By the rule we see that this will be an upward force of bCH. Since the direction of the current in the bottom of the rectangle is opposite to that in A′B′, this portion of the circuit will be acted upon by a downward force bCH, which will produce equilibrium with the force exerted upon A′B′. Since the current in the vertical side

A′D′ is upwards, the force on this side is still lCH in the direction EF, while the force on the vertical side B′C′ is lCH in the direction GK. Since these two forces are now not only equal and opposite, but act in the same straight line, they are in equilibrium. Thus, in this position, the rectangle is in a state of stable equilibrium, and there is no force, due to the magnetic field, either tending to move it bodily or to turn it about any axis.

From the above investigations we see that the rectangle tends to set itself in such a position that the number of tubes of force which pass through it is a maximum, and that the direction of the lines of force due to the circuit is, inside the coil, the same as that of the lines of force of the field.

If from analogy with a magnet we call the face of a coil at which the lines of force leave the space included by the coil the north surface of the coil, then we may express the results to which we have been led as follows: A circuit in which a current is flowing tends to turn so that the number of tubes of force due to the field entering its south face is a maximum. Since the direction of the lines of force due to the coil in the position A′B′ run inside the coil in the same direction as the lines of force due to the field, we may also summarise the result as follows: A coil in which a current is flowing will tend to set itself so that the total number of lines of force which pass through it are a maximum; that is, so that the total induction through the coil is a maximum.

511*. Magnetic Shell.—A thin plate of magnetic material which is magnetised so that the direction of magnetisation is everywhere perpendicular to the surfaces of the sheet is called a magnetic shell. The product of the thickness of the shell into the intensity of magnetisation is called the strength of the shell.

It can be shown that the magnetic force exerted by a closed circuit in which a current C is flowing is the same as that which would be exerted by a magnetic shell which occupied the space bounded by the circuit, and of which the strength was equal to C. The side of the shell which is a north pole must, of course, correspond to what was called the north side of the circuit in the last section. If S is the area of the space enclosed by the circuit, so that S is the area of the pole of the shell, the strength of each pole will be IS, for, as we have seen in § 501, the surface density of the free magnetism on the pole is numerically equal to the intensity of magnetisation. Hence the number of tubes of force which leave the north face of the shell and enter the south face, completing their course through the substance of the shell, is

$$4\pi SI.$$

The number of tubes of force which pass through any given area is called the induction through that area; or, to distinguish between this use of the term and that in § 502, we may call the total number of lines

of force passing through any given area the total induction through the area.

The reason why the magnetic effect of a shell is the same as that of a current flowing in a wire which has the form of the perimeter of the shell is at once apparent if we adopt Ampère's hypothesis as to the magnetism of the molecular magnets being due to currents circulating within the molecule. If, in Fig. 493, ABCD represents one face of the shell, then if the molecular amperian currents are represented by the small rectangles, the direction of the currents being indicated by the arrows, it will be seen that the molecular currents in the adjacent sides of any two contiguous molecules are in opposite directions, and hence they neutralise each other's effects as far as producing any external field is concerned. It is only in the molecules which bound the area ABCD that the molecular currents are able to exert any external force. But the combined effect of the currents in the bounding molecules will be the same as that of a current flowing round a wire which occupies the position of the edge of the shell.

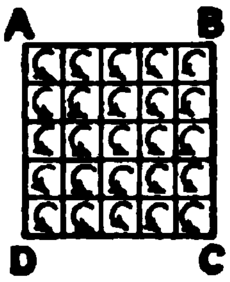


FIG. 493.

512*. Magnetic Moment of a Circuit conveying a Current.—Suppose that a conductor conveying a current C forms a plane closed circuit which encloses an area S. Then the magnetic effects of the circuit can be represented by a shell of which the strength is F, and which fills the space enclosed by the circuit. If the thickness of the shell is e, and the intensity of magnetisation is I, then $eI=F$. The magnetic moment of the shell is equal to the product of its volume into the intensity of magnetisation, that is, $M=eS.I$. But $C=F=eI$, hence

$$M=SC.$$

That is, the magnetic moment of the shell, which is equivalent to the circuit, is equal to the product of the strength of the current into the surface included by the circuit.

We may apply this result to obtain the turning moment acting on the rectangle ABCD in Fig. 492. Here the area of the circuit is ab, and hence the magnetic moment of the equivalent shell is abC. The axis of the magnet being at right angles to the plane of the coil, it is at right angles to the lines of force of the field. Hence by the formula obtained in § 425 the turning moment acting on the magnet is MH or $abC.H$, which agrees with the result already obtained from the consideration of the forces which act on the different portions of the circuit.

If, instead of being parallel to the lines of force of the field, the circuit makes an angle θ with the lines of force, then the axis of the equivalent magnetic shell will make an angle $90^\circ-\theta$ with the direction of the field, and the turning couple will be equal to $MH \sin(90^\circ-\theta)=MH\cos\theta$ $=SCH\cos\theta$.

If the circuit consists of a circular coil of n turns of insulated wire, the

radius of each turn being r, the moment of the equivalent shell will be $\pi r^2 nC$, for the area included by each turn is πr^2, and so the area included by the n turns is $\pi r^2 n$.

513*. Magnetic Field inside a Solenoid.—A cylinder which is lapped over with insulated wire, so that the circuit consists of a number of equal circles with their planes at right angles to the axis of the cylinder, is called a solenoid. If, as is generally the case in practice, the wire is coiled in a spiral, then it is equivalent to a number of circles placed in planes perpendicular to the axis and to a straight conductor which passes along the axis, for in the helix not only does the wire go round the cylinder, but it also is taken along parallel to the axis. If the solenoid consists of two layers of wire wound one on the top of the other, the direction of winding being the same in the two layers, and the wire all in one length, the longitudinal portion of the wire in the two layers will compensate the one for the other.

In order to calculate the strength of the magnetic field within a solenoid, the simplest way is to make use of the results obtained in § 511. Suppose that the solenoid consists of N turns, the radius of each turn being r, and that the length of the solenoid is L, so that there are N/L turns per unit of length. Calling the number of turns in the unit of length n, the length along the axis of the solenoid occupied by each turn is $1/n$. If the current flowing through the solenoid is C, then each turn can be replaced by a magnetic shell of which the perimeter is a circle of radius r, and the strength is numerically equal to C. If the thickness of each shell is taken as equal to $1/n$, the intensity of magnetisation will be Cn, for the strength of the shell is equal to the product of the thickness into the intensity of magnetisation. Since the effect of each turn of the solenoid can be represented by the effect of such a shell, the effect of the whole solenoid will be represented by the combined effect of N such shells placed end to end. Since the thickness of each shell has been chosen as equal to the distance between two adjacent turns of the wire, the shells when placed end to end will just occupy the space within the solenoid. The north and south faces of adjacent shells, being equally magnetised, will exactly neutralise each other's effect, so that the external effect will be due to the extreme faces only. Hence the magnetic effect of the solenoid will be the same as that of a cylindrical magnet of length L, and of which the cross-section is πr^2, when magnetised to the intensity I or nC. The total induction through such a magnet will be equal to $4\pi I \cdot \pi r^2$ less the demagnetising field produced by the poles at the ends. If the length of the solenoid is very great compared to its diameter, we may neglect the effects due to the ends, and the total induction through any cross-section near the middle will be $4\pi^2 r^2 I$ or $4\pi^2 r^2 nC$. The induction, that is, the number of tubes of induction per unit area of cross-section, is $4\pi nC$. Hence as the tubes of force, &c., are the same in the case

of the solenoid as in the case of the magnet we have been considering, we see that if the solenoid is very long compared to its diameter, the number of tubes of force which will cross any cross-section taken near the middle will be $4\pi nC \cdot \pi r^2$, or the number of tubes that cross unit area will be $4\pi nC$. But the number of tubes of force which cross unit area taken at right angles to the direction of the tubes is numerically equal to the strength of the magnetic field at the point. Hence the strength of the field near the middle of such a long solenoid is equal to $4\pi nC$, and so long as we are only considering a portion of the solenoid at some distance from either end the field will be uniform.

We have hitherto supposed that the interior of the solenoid is filled with a non-magnetic material, such as air. If we suppose that the solenoid is filled with a magnetic material such as iron, then if, as before, we suppose that the length is so great that the effects of the ends can be neglected, the magnetising force acting on the iron is $4\pi nC$. Hence if μ is the permeability of the iron for a magnetising field of this strength, the induction, B, through the iron will be $B=\mu H=4\pi\mu nC$. Thus the effect of the introduction of the iron is to increase the total induction through the solenoid from $4\pi^2 r^2 nC$ to $4\pi^2 r^2 nC \cdot \mu$.

When the solenoid was filled with air, we saw that its action could be represented by a magnet of the same length, of which the strength of each pole was $\pi r^2 nC$. This expression will not be exactly true, for some of the tubes of force will leave the solenoid before the end, passing through the sides; it is, however, sufficiently near the truth for the consideration of the general magnetic effect of the solenoid. When the solenoid is filled with iron, the number of tubes of induction leaving the iron near the north end of the solenoid will be $\pi r^2 B$ or $4\pi^2 r^2 nC\mu$, and the strength of the pole at each end will be $\pi r^2 nC\mu$. Thus by the introduction of the iron the moment of the magnet which would be equivalent to the solenoid is increased in the ratio of $\pi r^2 nC\mu$ to $\pi r^2 nC$ or as $\mu : 1$.

Since the value of μ may be as high as 2000, it is evident to what a great extent the magnetic effects of the solenoid are increased by the introduction of the iron.

Although the magnetic effect of a solenoid can be represented by a magnet, yet there is an important difference between a solenoid and a permanent magnet, in that in the case of a solenoid we have to do not only with the magnetic field in the space outside the magnet, but also with the space inside, the direction of the lines of force being opposite in the two regions. In the case of a permanent magnet, on the other hand, we are unable to utilise the magnetic field within the magnet. It might at first appear that the fields of force due to a solenoid and to a *hollow* cylindrical magnet would be of the same nature. This is however not the case, for inside the solenoid the lines of force run in the opposite direction to what they do in the space outside; while in the case of the hollow magnet the lines of force in the hollow run from the north pole to

the south, just as they do in the space outside. It is only inside the steel of the magnet that the lines of force run in the opposite direction to that in the space outside, namely, from the south pole to the north.

514. Action of Currents on Currents.—We have already considered the force exerted on a neighbouring magnet by conductors of certain simple shapes when they convey a current, or, in other words, the strength of the field produced by these conductors. We have now to consider the force exerted by one conductor conveying a current on another.

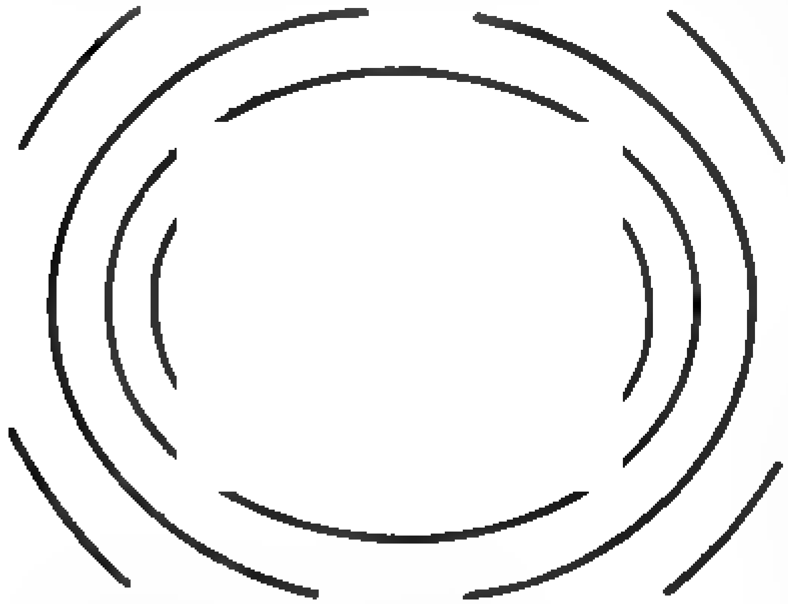


Fig. 494.

First consider the case of two long parallel conductors in which the currents are flowing in the same direction. Ampère, who first made experiments on the action of currents on currents, showed that under these circumstances the two conductors are attracted the one towards the other. A consideration of Fig. 494, which represents the lines of force of the two conductors, which are supposed to be perpendicular to the paper at the points A and B, will at once indicate that attraction must take place. For if we suppose that there exists a tension along the lines of force and a pressure at right angles, the lines of force which surround the two wires will, by their tension, tend to contract and thus to force the wires together.

Fig. 495.

In the case (Fig. 495) where the currents are flowing in opposite directions, the wires will be repelled, as was also shown by Ampère to be the case.

The magnitude of the force exerted on unit length of one of the conductors due to the other can at once be obtained. Let the current in the conductor which passes through A be C_1, and that in the conductor passing through B be C_2, both currents flowing upwards, the distance between A and B being d. Then the strength of the field at A due to the infinite conductor passing through B is by § 475 $2C_2/d$, and acts through A and in the direction of the arrow. Then we have a conductor passing through A in which a current C_1 is flowing at right angles to the direction of a magnetic field of strength $2C_2/d$; hence it experiences per unit of length a force of $2C_1C_2/d$ (§ 509), in a direction perpendicular to both the direction of the field and to that of the conductor, that is, parallel to the line AB. A consideration of the rule given in § 509 will show that the force acts in the direction $\overrightarrow{AB}$. If the direction of one of the currents is reversed the magnitude of the force will remain the same, but it will now act in the direction $\overrightarrow{BA}$, that is, it will be a repulsion. Since action and reaction are equal and opposite, the force on the conductor B will be in every case exactly equal and opposite to the force acting on the conductor A.

The attraction exerted by two parallel conductors, in which the currents are flowing in the same direction, can be very strikingly shown by means of a spiral of copper wire which is suspended from a stand by one end, and is allowed to hang free in such a way that the other end of the wire dips in a mercury-cup. If a current is sent through the spiral by connecting the fixed end to one pole of a battery and the other pole to the mercury, since the different turns of the spiral are parallel, and the current is flowing in the same direction in each, they will attract one another. As a result the spiral will contract, and in contracting the free end will be lifted out of the mercury and thus break the circuit and stop the current. As soon as the current is broken the attraction between the spirals will cease, and the spiral will elongate under the influence of gravity, and the end again coming in contact with the mercury the current will be started, and the whole cycle of operations again be gone through. Not only does this illustrate the attraction of parallel currents, but the apparatus is a machine by means of which electrical energy is continuously converted into mechanical energy, that is, into the energy of motion of the spiral.

515. The Electro-Dynamometer—Electric Balance.—The force exerted by one circuit in which a current is flowing on another circuit in which the same current is flowing is made use of to measure the strength of the current, and instruments for measuring currents depending on this principle are called electro-dynamometers.

One form of electro-dynamometer consists of two coils (A and B, Fig. 496) of large radius, placed with their planes parallel to one another and with their centres at a distance apart equal to the radius of either,

thus forming a Helmholtz galvanometer (§ 478). The current to be measured is sent round these two coils in the same direction, and also through a small suspended coil C, which hangs in a symmetrical position with respect to the coils, and has its plane vertical. The small coil is suspended by a bifilar suspension, D (§ 119), so that it is free to turn about a vertical axis, the wires of the bifilar suspension being utilised to lead the current into the coil. The suspending wires are so arranged that, when no current is passing through the coils, the plane of the suspended coil is at right angles to the planes of the large coils. When the current is passed through the fixed and suspended coils in series, the plane of the small coil being at right angles to that of the direction of the field due to the large coils, there will be a turning couple acting on the suspended coil which will tend to turn it into the same plane as the large coils. The bifilar suspension will oppose this turning, and the suspended coil will take up a position such that the deflecting couple due to the electromagnetic forces will be equal and opposite to the restoring couple due to the bifilar. The magnitude of the deflecting couple can be calculated in terms of the current by the processes we have already given, for the strength of the field due to the large coils at the place where the suspended coil is hung can be calculated by the method given in § 476, if the radius of the circles and the number of turns of the wire in each is known, while the couple acting on the coil C, when its axis makes any given angle with the direction of this field, can be calculated by the process given in § 512. The value of the couple due to the bifilar can also be calculated from the weight of the

Fig. 496.

coil and the length and distance apart of the suspending wires. Hence, equating these two couples, the value of the current, which is the only unknown quantity, can be calculated.

In performing the experiment, account must be taken of the earth's field. This, however, may be eliminated if the whole instrument is turned so that when the suspended coil comes to rest under the action of the current its axis lies magnetic north and south.

Another method of determining the absolute value of a current consists in measuring the vertical force exerted on a small horizontal coil A (Fig. 497), which is attached to one of the arms of a balance, owing to the magnetic action of two fixed coils, C and D. The current to be measured is passed through the fixed coils and the suspended coil in series, in such a way that the suspended coil is repelled by one of the fixed coils and attracted by the other. The force with which the fixed

FIG. 497.

coils act on the suspended coil can be calculated in terms of the current which is passing from the dimensions of the coils, and hence, by measuring this force by means of the weights which have to be put in the scale pan E of the balance in order to bring the beam into the horizontal position, the value of the current can be calculated.

Using this method, Lord Rayleigh found that the current, which by the balance was equal to 1 ampere, when passed through an electrolytic cell (§ 539) containing a solution of silver nitrate with a silver anode and a platinum kathode, deposited 0.0011179 grams of silver in one second.

This value for the electro-chemical equivalent (see § 539) of silver is generally made use of to standardise instruments for the measurement of currents in which the value of the current corresponding to a given indication cannot be directly calculated from the dimensions of the instrument.

PART VII.—ELECTRO-MAGNETIC INDUCTION

CHAPTER XIII

INDUCED CURRENTS

516. Induced Currents.—In 1831 Faraday took a ring of iron and on it wound two coils of insulated wire. Having connected one of these coils to the terminals of a galvanometer, he passed an electric current through the other coil, and then found that at the moment of starting the current the needle of the galvanometer was deflected, showing that a current was passing in the second closed circuit. This deflection was only momentary, and the galvanometer immediately came back to its undeflected position, although the current in the magnetising coil was still flowing. On breaking the current, however, another momentary deflection of the galvanometer took place, but in the opposite direction to that which had occurred when the current was started.

He next wound two coils alongside one another on a wooden cylinder, and again found that when an electric current was either started or stopped in the one coil, a galvanometer connected with the other coil indicated the passage of a momentary current, the direction of the current when the main current was started being in the opposite direction to that obtained when the current was stopped.

Finally he found that if a magnet is inserted into a coil, at the instant when the magnet is inserted a current is produced in the coil, and that when the magnet is withdrawn a current in the opposite direction is also produced. He also found that if a wire, the ends of which are connected to the terminals of a galvanometer, was passed between the poles of a powerful horse-shoe magnet, so that the direction of motion of the wire was such that it cut across the lines of force of the magnet, then a current was produced during the time that the wire was being moved across the lines of force.

The currents which are produced in these ways in a closed circuit when a current in a neighbouring circuit is started or stopped, or by the relative motion of the circuit and a magnet, are called *induced currents*.

These results obtained by Faraday, which, as we shall see, are the foundation on which are based all the modern methods of producing the

currents that are used in such numberless ways, such as in the production of light, the moving of vehicles, driving machinery, and performing many chemical processes, can all be summed up in the following short law :—

Whenever, from any cause whatever, the number of tubes of force which thread through any conducting circuit is altered, an electromotive force will be produced during the change in the number of tubes which will produce or tend to produce a current in the circuit.

517. Lenz's Law.—The direction in which the induced currents flow has been put into a concise form by Lenz, in what is known as Lenz's law, and is as follows :—

The direction of the induced current produced in a conductor due to the movement of a magnet, or to that of a circuit in which a current is flowing, is always such as, by the action of the induced current on the magnet or current-conveying conductor, to produce a force tending to oppose the motion.

Thus suppose there are two parallel conductors, in one of which a current is flowing, and that the distance between the conductor is decreased, then the direction of the induced current will be such as to oppose the motion, that is, will be such as to cause repulsion between the conductors. Hence, since repulsion takes place when the currents are in opposite directions, it follows from Lenz's law that, when the conductors are moved nearer, the induced current will be in the opposite direction to the inducing current.

If, instead of the distance between the conductors being altered, the current is either started or its strength is increased, then we may look upon this as being the same thing as if the conductor in which the current is started were moved up to its present position from an infinite distance. While if the current is stopped or its strength decreased, then this is equivalent to the conductor in which it flows being removed from the neighbourhood of the conductor in which the electro-magnetic induction is produced. Hence it follows from Lenz's law that the direction of the induced current when the current is started or increased in strength is the same as when the conductors are moved nearer together, namely, in the opposite direction to that in the inducing current ; while when the current is stopped the direction of the induced current is the same as that of the inducing current before it was stopped.

518. Electro-magnetic Induction.—We will now proceed to examine different cases of the production of induced currents from Faraday's point of view as to tubes of force.

In the first place, the phenomenon of the production of induced currents is said to be due to electro-magnetic induction. It is called electro-magnetic induction rather than, as is sometimes done, simply induction, for the sake of preventing confusion with the use of the term induction given in § 502.

The conductor in which the inducing current flows is called the

primary conductor, or simply the primary, while the conductor in which the induced current is produced, or at any rate in which an induced electromotive force is developed, is called the secondary.

First let us consider the case of a primary which consists of a single circle of wire, P (Fig. 498), in which a current is caused to flow by a battery, B, and in which there is a key, K, by means of which the circuit can be closed or opened, and thus the current started or stopped. Let the secondary, S, consist of a similar circle of wire, in which, if we like, we may suppose a galvanometer, G, is included. If a current in the direction of the arrow is flowing in the primary circuit, tubes of magnetic induction will thread through the primary in the direction shown, and some of these will also thread through the neighbouring secondary. Suppose now that the current in the primary is stopped, then all the tubes of induction due to this current will vanish. Hence the number of tubes of induction which thread through the secondary will be diminished, so that an induced current will be produced. From Lenz's law it follows that the direction of this induced current will be the same as that of the current in the primary. Now the induced current in the secondary will produce tubes of induction, and since the direction of the induced current is the same as that of the primary current, the direction of the tubes of induction due to the induced current will be the same as those due to the primary current. Some of these tubes will thread through the primary circuit, so that the effect of the induced current is to tend to keep the number of tubes which thread through the primary circuit constant, although on account of the stoppage of the primary current the number of tubes tends to become less. The same effect occurs in the secondary circuit, for the tubes due to the induced current, which are introduced when the primary current is broken, are such as to tend to keep the induction through the secondary constant.

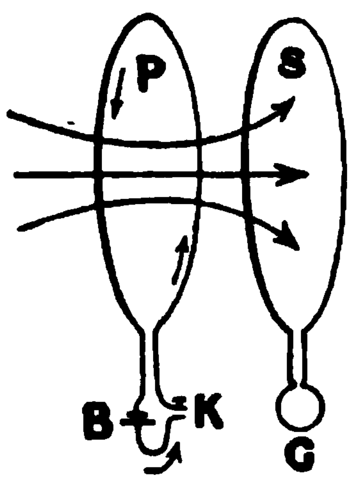


FIG. 498.

Next take the case where the current in the primary is started. The direction of the induced current is opposite to that of the primary current, hence the tubes of induction which thread through the secondary, due to the induced current, are in the opposite direction to those which are being threaded through the circuit due to the starting of the primary current. Hence the total induction through the secondary during the time the induced current lasts is the difference of the induction through this circuit due to the primary and the induced currents, so that in this case also the induced current is such that it tends to keep the total induction through the secondary circuit the same as it was before the starting of the primary current. Also, since some of the tubes due to the induced current will thread through the primary, the effect of the presence of the secondary will be to postpone the time when the number of tubes of induction

through the primary reaches its final value, since their presence tends to keep the induction through the primary the same as it was before the starting of the current.

In the case when the current in the primary is kept constant, but the distance between the primary and secondary circuits is varied, the same effect takes place, namely, the induced currents are in such a direction as, by their action, to keep the number of tubes of induction which pass through the secondary circuit the same as it was before the motion. Of course, since the induced current only lasts while the number of tubes of induction is varying, the *final* induction through the secondary, as well as that through the primary itself, is quite unaltered by the fact that an induced current is produced.

Although when the secondary conductor does not form a closed circuit no induced current will flow in the secondary, yet in this case there will be an electromotive force produced owing to the electro-magnetic induction, the direction of the E.M.F. being such that if the circuit were closed the current which would be produced by this E.M.F. would be that which we have been considering in the case of a closed secondary.

In the case of an unclosed secondary circuit, since there will be no induced current, there will be no tubes of induction due to the induced current, which, by being threaded through the primary circuit, will tend to delay the induction through this circuit from at once attaining its final value. In this case, as well as in that where there is no secondary near a circuit in which a current is started or stopped, we might expect that the current would instantly attain its final value when the circuit is closed. This, however, is not the case, for the circuit itself acts in such a way as to tend to keep the induction through itself constant. Thus before the current is started there are no tubes of induction passing through the circuit, but when the current is passing there are a certain number of these tubes. Hence the number of tubes of induction threading through the circuit has been increased, and during the time that they were being threaded through there will be an induced current produced in the circuit itself, just as in the case of a circuit in which the increase of the total induction is due to some other circuit. As we have seen, the direction of the induced current is such as to tend to keep the number of tubes of induction linked through it constant. Hence when the current is started, so that the number of tubes is increased, the induced tubes must be in the opposite direction to those due to the current which is being started, that is, the direction of the induced current must be the opposite to that of the current which is started. The effect of this induced current, which is said to be due to *self-induction*, is to delay the current in the circuit attaining its full value, though it has no effect on the final value which the current will reach; which final value of the current, in a simple metallic circuit, is that given by Ohm's law.

When the current in a circuit is stopped, the induced current, in order

that it may tend to keep the induction through the circuit constant, must be in the same direction as the main current. The presence of the induced current when a current is stopped can be very clearly shown by means of the arrangement shown in Fig. 499. A coil, C, which ought to have a large number of turns, and an iron core (it will be remembered how the presence of an iron core increases the induction through a coil) is connected up with a make and break key, K, and a battery, B. An incandescent electric lamp, L, is connected in parallel with the coil. Although the battery may not be of sufficiently high electromotive force to cause the lamp to glow when the current is passing round the circuit, yet, when the key is opened, the induced E.M.F., due to the self-induction of the coil, will be so great that sufficient current will flow through the coil and the lamp circuit to cause the lamp to glow brightly.

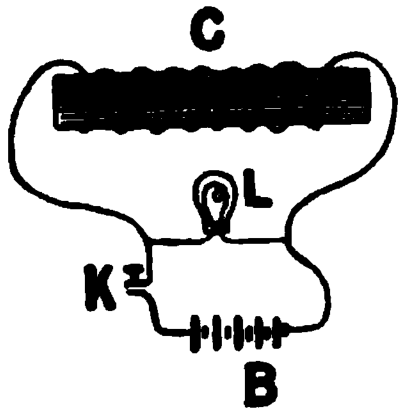


FIG. 499.

519. Magnitude of the Induced E.M.F.—We have hitherto only considered the conditions under which induced currents are produced and the direction in which they flow, and we now have to consider on what conditions the magnitude of the induced current depends.

In the first place, the magnitude of the induced current depends on the resistance of the secondary, and since, other things remaining the same, the current is inversely proportional to the resistance, we shall in future consider the electromotive force induced in the circuit considered on account of induction; and where the value of the induced current is required, this can be calculated according to Ohm's law. The expression for the magnitude of the induced electromotive force was first given by Newmann. We may combine Newmann's results with Faraday's law as to electro-magnetic induction as follows:—

Whenever the number of tubes of induction which thread through a circuit is altering, an E.M.F. is produced in the circuit numerically equal to the rate at which the number of tubes of induction is diminishing.

The direction of the E.M.F. obtained by this rule is positive when it tends to produce a current in the circuit which is related to the direction of the lines of force, as the direction of rotation of a corkscrew is related to the direction of translation. Or if we are looking along the tubes of induction towards the circuit, then, if the number of tubes is decreasing, the E.M.F. will act in the clockwise direction round the circuit.

The direction of the induced E.M.F. in a straight conductor, which is moving in a direction at right angles to the lines of force of a magnetic field, can be remembered by the following rule:—

Hold your *right* hand with the fingers pointing towards the direction in which the conductor is moving, and with the palm turned in the direction in which a north pole would travel in the field, *i.e.* so that the tubes

of force enter the hand at the back, then the outstretched thumb will give the direction of the induced E.M.F. in the conductor.

In a uniform field of strength F the cross-section of the tubes of induction, or, what is the same thing, the tubes of force, is $1/F$, so that F tubes cross unit area at right angles to the direction of the tubes. If then a straight conductor of length L is moved with a velocity v in a direction perpendicular both to the tubes of induction of the field and to the length of the wire, the space swept out by the wire in unit time will be vL. Hence the number of tubes of force cut through by the wire in unit time will be $vL.F$. This then is the rate at which a circuit, of which the conductor forms a part, is increasing the number of tubes of induction which it embraces. Hence the electromotive force induced in the conductor due to the cutting of the tubes of induction of the field is vLF.

520. The Earth Inductor.—Suppose that during a very small time δt the total induction through a circuit is changed from B to $B-\delta B$. Then the number of tubes of induction which thread through the circuit have in the time decreased by the amount δB, so that the induced E.M.F. acting round the circuit, which is equal to the rate at which the number of tubes is decreasing, will be equal to $\delta B/\delta t$. Hence, if the resistance of the circuit is R, the induced current will be $\delta B/R\delta t$. If the rate at which the tubes of induction leave the circuit remains constant during the time δt, the current induced will also be constant. Hence the quantity of electricity which will flow round the circuit in the time δt will be $C\delta t$ or $\delta B/R$. Thus when the rate at which the tubes of induction are cut by the circuit is constant, the quantity of electricity which will flow round the circuit in any given time will be equal to the quotient of the number of tubes cut by the circuit in this time by the resistance of the circuit. The same relation must also hold if the rate at which the tubes are cut is not uniform, for we can divide the interval considered into a number of very small intervals each equal to δt, in each of which the rate at which the tubes are being cut is sensibly uniform, and the quantity of electricity which traverses the circuit during each of these intervals is given by the above expression. Hence, adding together all the intervals, we get on one side the total quantity of electricity which passes during the interval considered, and on the other side the sum of the decreases of the number of tubes of induction, that is, the total decrease in the number of tubes divided by the resistance of the circuit.

If a circuit, which encloses an area A, is placed in a uniform magnetic field of strength H with its plane at right angles to the direction of the field, the total number of tubes of induction passing through the circuit will be AH, for the cross-section of a tube is $1/H$, hence H tubes cross unit area taken perpendicular to the direction of the tubes, and AH will cross an area A. If the circuit is now turned till its plane is parallel

to the direction of the tubes of force of the field, none of these tubes will pass through the circuit, and hence the total induction through the circuit will be zero. Thus by rotating the circuit in this way the number of tubes of induction passing through the circuit has been decreased by AH, and hence, if R is the resistance of the circuit, the quantity of electricity which passes round the circuit due to induction will be AH/R. The direction in which this electricity will be displaced round the circuit will be the clockwise direction as we face the side of the circuit at which the tubes of induction enter the space enclosed by the circuit.

If now the circuit is rotated through another right angle in the same direction as before, so that it is again at right angles to the direction of the field, but with the tubes of induction threading through in the opposite direction to what they did before, the number of tubes which thread through will during the rotation increase from zero to AH. The result will be that a quantity of electricity equal to AH/R will again be caused to circulate round the circuit, and since the tubes are now inserted instead of being withdrawn, the direction in which the current will circulate will be the anticlockwise one if, as before, we face the side where the tubes enter the circuit. Since, however, the position of the circuit is reversed from what it was in the former case, the direction of the induced current, as far as the circuit itself is concerned, will be the same as before. If, then, starting with the circuit perpendicular to the lines of force of the field, it is rotated through 180°, the quantity of electricity which will traverse the circuit due to the electro-magnetic induction will be given by

$$Q = 2AH/R.$$

Hence, if we measure Q by passing the induced current through a galvanometer, in which case the quantity R must include the resistance of the coil with the connecting wires and the galvanometer, and we know the quantity A, we can calculate the strength of the magnetic field in which the coil is rotated.

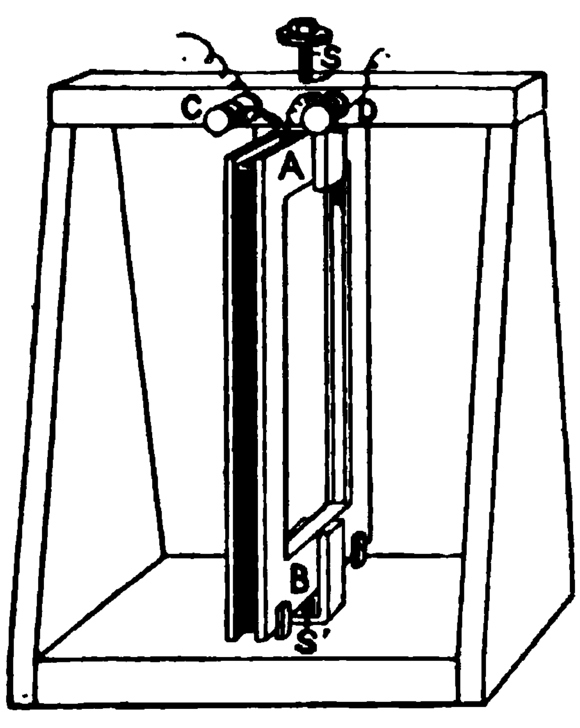


FIG. 500.

An instrument for measuring the strength of magnetic fields on this principle is shown in Fig. 500, and is called an earth inductor, from the fact that it is often used to measure the strength of the earth's field. A coil, AB, containing a number of turns of insulated wire, and for which the area included by each turn has been measured so that the quantity A, which is the sum of the areas included by all the turns, is known, is supported in a frame in such a way that it can be rotated through two right angles about an axis SS′. The ends of

the coil are connected to a galvanometer by means of which the quantity of electricity which passes round the circuit can be measured. If the axis SS′ is placed vertical, the plane of the coil being set magnetic east and west, and the coil is rotated through two right angles, and if Q is the quantity of electricity which passes through the galvanometer, we have $Q=2AH/R$; for, as the coil has been turned about a vertical axis, it has not intersected any of the tubes of induction of the vertical component of the earth's field, and hence the only part of the field with which we are concerned is the horizontal component, H.

Next suppose the axis of rotation of the coil to be horizontal and in the magnetic meridian, and that, starting with the coil in a horizontal plane, it is again rotated through two right angles, the quantity of electricity, Q', induced being measured. In this case, since the axis about which the coil is rotated is parallel to the direction of the horizontal component, the coil will not cut through any of the tubes of induction corresponding to the horizontal component. Hence we have $Q'=2AV/R$, where V is the vertical component of the earth's field. Dividing one of these expressions by the other, we get

$$V/H=Q'/Q.$$

But if θ is the dip, then $\tan\theta=V/H$. Hence, from the result of two observations with the earth inductor, one with the axis of rotation vertical, and the other with the axis horizontal, we can calculate the value of the dip at the place where the observations are made.

521. Determination of the Value of the Ohm by the B.A. Committee.—The first measurement of a resistance in absolute measure of any accuracy was performed by a committee of the British Association, by a method involving the induced E.M.F produced in a coil when rotated in a magnetic field. The method they employed consisted in spinning a coil, of which the two ends of the wire were joined together, about a vertical axis in the earth's field, and noting the deflection produced by the currents induced in the coil on a magnetic needle placed at the centre of the coil. If the dimensions of the coil are known, and they were determined during the winding, and the speed of rotation when the deflection is measured is known, it is possible to calculate the resistance of the coil in absolute measure, that is, in *c.g.s.* units. For, as we shall see in § 526, when a coil rotates in a magnetic field an induced E.M.F. will be produced, and the current which this E.M.F. will send round the coil will be inversely proportional to the resistance of the coil, while the deflection of the needle will be proportional to the current, that is, inversely proportional to the resistance.

From the results of their measurements the Committee constructed a number of coils consisting of platinum silver alloy, the resistance of each of which was equal to what from their measurements appeared to be the value of the ohm as defined in § 480. These coils have been preserved,

and are known as the B.A. units. More recent measurements made by the same method, as well as by several different methods, have shown that the B.A. units are not exactly 1 ohm, the true value being, 1 B.A. unit = 0.9866 ohm.

Experiment has also shown that the resistance of a column of pure mercury 106.3 cm. long and one square millimetre in cross-section, when at a temperature of 0° C., is equal to one ohm. As the resistance of a solid is dependent on the physical state, such as the hardness, &c., there is some doubt whether a standard resistance composed of a wire may not alter in time, due to a change in the molecular state of the metal. In the case of a liquid, however, such a molecular change is not to be feared, for liquids are not able to take up a state of strain. For this reason the final standards of resistance are composed of tubes of glass, of which the dimensions can be accurately measured, filled with pure mercury, and the wire standards used in ordinary work are compared with these mercury standards.

522. Determination of the Value of the Volt.—If the absolute value of a current and of a resistance is known from the measurements made by the methods described in §§ 515, 521, then, by passing the current through the resistance, the absolute value of the E.M.F. between the terminals of the resistance will be known. Hence of the three electrical quantities, ampere, ohm, and volt, the knowledge of the absolute value of any two enables us to obtain, by means of Ohm's law, the value of the third.

The values for the E.M.F. of the Clark and cadmium standard cells, given in §§ 554, 555, have been determined by comparing their E.M.F. with that developed between the terminals of a wire of known resistance when a current is passed, the value of the current being obtained from the indications of a current balance.

523. Arago's Experiment—Foucault Currents.—Arago discovered that if a copper disc is rotated about a vertical axis below a pivoted magnetic needle, the needle is deflected in the direction in which the disc is rotating, and if the speed of rotation is fairly great the needle is dragged completely round, so that it is set in rotation. The inverse experiment can also be performed, that is, if a magnet is rotated near the face of a copper disc which is free to turn, the disc is set in rotation, the direction of rotation being the same as that in which the magnet is being rotated. The explanation of this experiment was given by Faraday, who showed that it was due to the reaction between the electric currents induced in the copper disc and the magnet.

Let AB (Fig. 501) be the copper disc which is rotated in the direction shown by the arrow, and let NS be a magnet suspended or pivoted above the surface of the disc. The tubes of induction of the magnet pass from the north pole N to the south pole S, spreading out in the air. Some of these tubes will pass down below the copper disc near the pole N,

and will come up through the disc near S. Hence when the disc is set in rotation we have the portions of the conducting disc near N and S moving so as to cut through the tubes of induction, and hence an E.M.F. will be set up which will cause currents to circulate in the disc. Now by Lenz's law the direction of the induced currents must be such as to tend to check the motion, that is, such that the force which will be called into play between these induced currents and the inducing magnet will be so directed as by their action to check the motion which causes the induced currents. Hence, since action and reaction must always be equal and opposite, a force will act on the magnet tending to move it in the same direction as that in which the disc is rotating. The direction in which the currents must flow, so as to tend to turn the magnet in the same direction as that in which the disc is rotating, can be obtained by making use of the rule given in § 471. If we imagine a man in the disc near N facing the pole so that he must be on his back, then, in order that the pole N may be urged towards his left hand, he must lie with his head towards the circumference of the disc and his feet towards the centre. If when he is in this position a current is flowing from his feet to his head, the magnet pole N will be urged towards his left hand, that is, in the same direction as that in which the disc is rotating. In the same way it can be shown that, in order to produce a force between the magnet and the disc tending to check the motion of the disc, the currents in the portion of the disc near S must flow from the circumference towards the centre of the disc. Hence the path of the induced currents in the disc is somewhat as shown by the dotted curves.

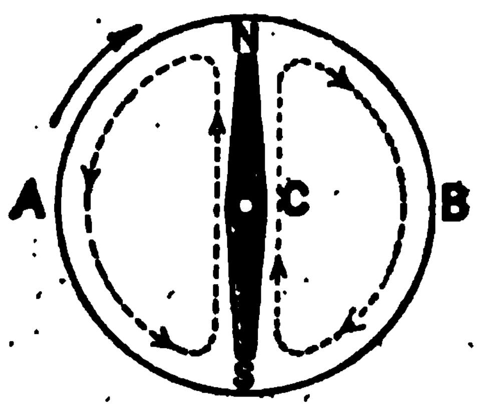


FIG. 501.

Of course, we could have arrived directly at the same result by making use of the rule given in § 519 for the connection between the direction in which a conductor is moved through a magnetic field and the direction of the induced E.M.F. Thus the portion of the disc under the pole N is moving in a magnetic field where the tubes of induction are running downwards, and hence, if the right hand is placed palm downwards with the fingers pointing in the direction the portion of the disc below N is moving, the outstretched thumb will give the direction of the induced E.M.F., and this will act from the centre of the disc towards the circumference.

The effects of the currents induced in a mass of metal when it is moved in a magnetic field was very strikingly shown by Foucault, who arranged a copper disc so that it could be rotated by means of a handle and a train of wheels between the poles of a powerful electro-magnet. Although it was easy, when the magnet was not excited, to rotate the disc at a rapid rate, on starting the current in the electro-magnet

the reaction on account of the induced currents was so enormous that the disc was immediately brought almost to rest, and it could only be rotated at a comparatively slow speed. These currents, which are induced within a mass of metal when it is in a changing magnetic field or is in motion in a steady field, are generally called Foucault currents. The circulation of the currents is of course accompanied by the conversion of electrical energy into heat according to Joule's law, so that the mechanical energy which has to be spent in moving the conductor appears as heat developed in it.

Use is often made of Foucault currents to check the oscillations of a suspended magnetic needle, such as a galvanometer needle, which are often a source of considerable loss of time, since the needle takes some time in coming to rest after it has been deflected. If the needle is surrounded by a thick copper box made to fit as near the needle as possible, when the needle is in oscillation induced currents will be produced in the copper, which will tend to check the motion of the needle. Under these circumstances the motion of the needle is said to be damped.

524. The Induction Coil.—By means of electro-magnetic induction, it is possible to produce in a secondary circuit an induced E.M.F. which is higher than the E.M.F. employed to produce the current in the primary circuit. If, on account of the current passing in a primary circuit, n tubes of induction pass through a secondary which consists of a single turn, the induced E.M.F. produced when the current in the primary is varied is equal to the rate of change of n. If, however, the secondary circuit consists of two turns, so that the n tubes of force due to the primary thread through both turns, the E.M.F. induced in *each* turn will be equal to the rate of change of n, and hence the total E.M.F. produced in the circuit will be the sum of the E.M.F.'s produced in the two portions of the circuit, that is, will be equal to twice the rate of change of the number of tubes of induction which pass through the secondary. Thus, by increasing the number of turns of the secondary circuit, the induced E.M.F. produced by a given rate of change of n can be made very great. One of the best known arrangements for obtaining a very high E.M.F. by means of electro-magnetic induction is the induction coil which, since it was first employed by Ruhmkorff, is often called Ruhmkorff's coil. The primary of these coils consists of a comparatively few number of turns of fairly thick wire, which is wound on a core composed of soft iron wires. The object of the iron core is to increase the induction through the primary produced by any given current, as was explained in § 513. The reason why wires are used instead of a solid rod is to prevent, as much as possible, the formation of Foucault currents in the mass of the iron, since these currents would not only waste the electrical energy used to work the coil, but would also, by their reaction on the primary current, tend

to keep this current from changing rapidly. For they would produce tubes of induction in such a direction as to keep the total induction through the primary constant when the strength of the primary current is altered. The iron used must be of a very soft quality, so that the hysteresis and residual magnetism which it possesses may be as small as possible, for the effect of hysteresis is to convert some of the electrical energy into heat as well as to make the changes in the induction through the coil slower. Round the outside of the primary coil is wound a secondary coil consisting of a very large number of turns of fine wire, each turn being very well insulated by means of a covering of silk and shellac. The ends of the secondary are generally connected to two insulated brass rods, the ends of which form a spark-gap of adjustable length. The current in the primary circuit being alternately made and broken, the induction through the secondary changes and an induced E.M.F. is produced in the secondary, which is in one direction when the current is made and in the opposite direction when the current is broken. Various arrangements are employed for automatically making and breaking the primary current. In some of these a small electric motor makes and breaks the current by dipping a rod of platinum into a mercury-cup. The more usual arrangement, at any rate on small coils, is to have a small piece of iron fixed to the end of a spring, so that when the current passes and magnetises the iron core the piece of iron is attracted. When no current is passing, the spring keeps the iron away from the end of the core, and makes contact between a piece of platinum fixed to the back of the iron and a platinum point which is attached to a pillar carried by the base of the coil. The primary current passes between the platinum point and the spring, and hence when the iron hammer is attracted by the core the primary current is interrupted. The interruption of the current causes the core to lose its magnetism, so that it no longer attracts the hammer, and hence the spring forces it back against the platinum point, thus again completing the primary circuit.

Since the magnitude of the induced E.M.F. depends on the rate at which the number of tubes which thread through the secondary change, it is of importance to make the starting and stopping of the primary current as sudden as possible. Now it has been shown in § 518 that, on account of the self-induction of a circuit in which a current is stopped or started, the current does not reach its full value at once, nor does it die away instantaneously. The effect of self-induction is shown very markedly by the spark which is produced every time the primary current is broken. It has been found that the intensity of the spark formed at the break can be considerably decreased, and hence the rapidity with which the primary current stops increased, so that the induced E.M.F. is also increased, by using a condenser, formed by a number of sheets of tinfoil separated the one from the other by sheets of

paraffined paper, one armature being connected with the spring of the interrupter and the other with the platinum point. In this way the condenser and the primary coil are connected in parallel, and it can be shown that connecting a condenser in this way has the same effect as if the self-induction of the coil were reduced.

By means of such a coil it is possible to produce a spark between the terminals up to about 20 inches in length, and this when the E.M.F. used to produce the primary current is only a few volts, and would be quite unable to produce a spark of a hundredth of an inch in length. Although the E.M.F. of the induced current is very great, the quantity of electricity which traverses the secondary is excessively small, for, on account of the great length of the secondary wire and its small diameter, the resistance of the secondary is very great.

CHAPTER XIV

ELECTRO-MAGNETIC MACHINES

525. Barlow's Wheel.—One of the simplest arrangements for converting electrical energy into mechanical energy is that known as Barlow's wheel, and is shown in section in Fig. 502. A copper disc, A, is mounted on a horizontal axle, the bottom edge of the disc just dipping into some mercury placed in a small dish D. The disc A turns between the poles of a magnet, NS, and a current is passed through the disc between the mercury dish D and the axle. Thus in the portion of the disc A between the poles of the magnet we have an electric current flowing at right angles to the lines of force of the field, and therefore the conductor conveying the current, that is, the disc, is acted upon by a force tending to move it at right angles to the lines of force and to the direction of the current, *i.e.* to rotate the disc about the axle.

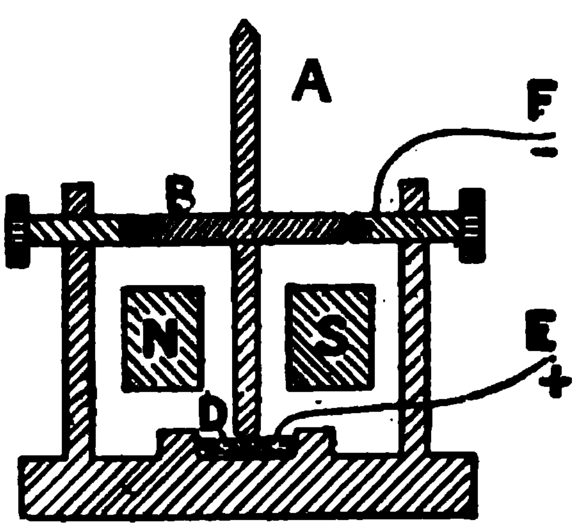


FIG. 502.

If the wheel is rotated by mechanical means, and the wires E and F are joined together, a current will be produced in this circuit, for the portion of the circuit which is formed by the radius of the disc between the axle and the mercury-cup will be moving at right angles to the lines of force of a magnetic field, and hence will be the seat of an induced E.M.F. The direction of the rotation of the wheel in the first case, and that of the induced current in the second, can easily be obtained by means of the rule given in § 519.

526. Induced Currents produced by Rotating a Coil in a Magnetic Field.—Suppose that a rectangular circuit of length a and breadth b is rotated about an axis AB, Fig. 503 (a), which is at right angles to the lines of force of a uniform field of strength H, and that the ends of the rectangle are connected with a stationary circuit, the resistance of this circuit and of the rectangle being R. Let us start with the rectangle in the position CD, Fig. 503 (b), in which it is at right angles to the lines of force of the field, so that the number of tubes passing through the rectangle is $ab.H$. Suppose now that the rectangle is turned into the position C′D′, making an angle θ with CD. The number of tubes which now pass

through the rectangle is evidently equal to the apparent area of the rectangle, as seen in the direction of the tubes, multiplied by H. But the area, as seen in the direction of the tubes, is equal to $a.\overline{EF}$ or $2a.\overline{EA}$. But $\overline{EA}=\overline{AC}\cos\theta=b/2.\cos\theta$. Hence the number of tubes of induction passing through the rectangle in its new position is $abH\cos\theta$. If the angular velocity of the coil is uniform and equal to ω, and if t is the time since the coil started from the position AB, we have $\theta=\omega t$. Now suppose that in the small time δt the coil turns through the angle $\delta\theta$. The number of tubes now passing through the circuit will be $abH\cos(\theta+\delta\theta)$. Hence in the small time δt the number of tubes has decreased by $abH\{\cos(\theta+\delta\theta)-\cos\theta\}$. Now $\cos(\theta+\delta\theta)=\cos\theta\cos\delta\theta-\sin\theta\sin\delta\theta$. If $\delta\theta$ is very small $\cos\delta\theta=1$ and $\sin\delta\theta=\delta\theta$, so that the decrease in the number of tubes in the time δt is $abH\sin\theta.\delta\theta$. Now the decrease in the number of tubes divided by the time during which this decrease takes place is, if the decrease goes on at a constant rate, and since δt is very small, we

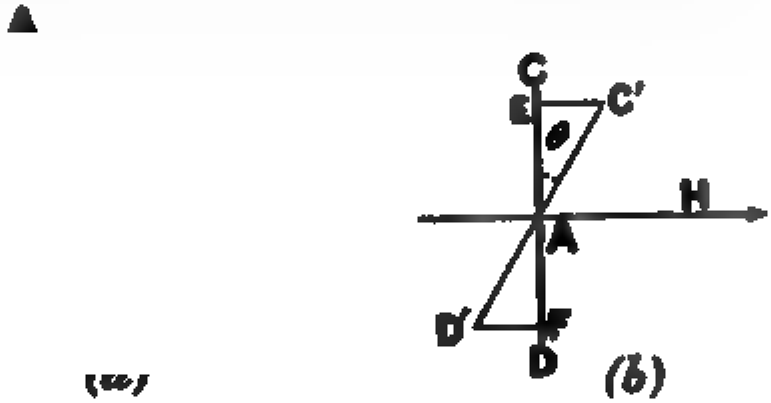


FIG. 503.

may consider that at any rate during this time that this is so, equal to the rate of decrease of the number of tubes, and, as we have seen, this is equal to the induced E.M.F. Hence the induced E.M.F. is equal to $abH\delta\theta\sin\theta/\delta t$, or $abH\omega\sin\theta$. Hence the induced electromotive force is at any time proportional to the sine of the angle which the plane of the coil makes with the lines of force of the field. Thus if E is the induced E.M.F. at a time t after the instant when the coil passed through the position CD, we have

$$E=SH\omega\sin\theta=SH\omega\sin\omega t,$$

where S has been written for the inductive area of the coil. Since the resistance of the coil and its connected circuit is R, an electromotive force E will produce a current C given by the relation $C=E/R$. Hence if the current in the coil at a time t is C, we have

$$C=\frac{SH\omega}{R}\sin\theta=\frac{SH\omega}{R}\sin\omega t.$$

Thus, as the coil rotates, a periodic current and E.M.F. will be pro-

duced, the maximum current occurring when $\theta=90°$, so that $\sin\theta=1$, the maximum value of the E.M.F. being ωSH, and that of the current $\omega SH/R$. When the plane of the coil is at right angles to the lines of force of the field $\theta=0$ or $180°$, and $\sin\theta=0$, so that the induced E.M.F. and also the current is zero. While θ changes from 180° to 270°, the induced E.M.F. changes from zero to $-\omega SH$, and the current increases from 0 to $-\omega SH/R$. The minus sign shows that the current is in the opposite direction to what it was while θ increased from 0° to 90°. For $\theta=270°$ the current is again a maximum, but, as we have pointed out, in the negative direction. As θ changes from 270° to 360°, the current decreases to zero, while as θ changes from 0° to 90° the current is again in the positive direction and increases from 0 to $\omega SH/.R$.

Thus in the circuit attached to the coil a current will be produced which changes its direction twice in each revolution of the coil, the maximum current in each direction being the same. Such a current is called an alternating current.

By suitable arrangements this alternating current in the circuit attached to the coil can be changed into a current which always flows in the same direction. Under these circumstances the alternating current is said to be rectified. A method of rectifying the current consists in fitting a copper ring on the axle on which the coil turns, which is insulated from the axle, and is in addition split along two generating lines which are on opposite sides of the ring as shown at *abcd*, Fig. 504. Two copper springs, B_1 and B_2, called brushes, rest against the copper ring, and are connected to the two ends of the external circuit. One end of the coil is connected to *ab* and the other to *cd*. The positions of the two brushes, B_1, B_2, are so arranged that as the coil revolves the brushes cross the gaps *ad* and *bc* in the ring, just as the coil is passing through the position in which its plane is perpendicular to the lines of force of the field, and hence the induced current is zero. Suppose that when the coil is in the position C′D′ (Fig. 503) the end of the coil connected with *ab* is at the higher potential, so that the current in the external circuit is going from B_1 to B_2. When the coil has passed through the position in which its plane is at right angles to the lines of force of the field, the direction of the induced E.M.F. will be reversed; thus *dc* will now be at the higher potential. The copper conductor *dc* will now be in contact with the brush B_1, and hence the current in the external circuit will still flow from B_1 to B_2. Although the current in the external circuit is now always in the same direction, it is not a constant current, but twice in every revolution it is zero, and twice reaches a maximum value of $\omega SH/R$. The difference between this rectified current and the

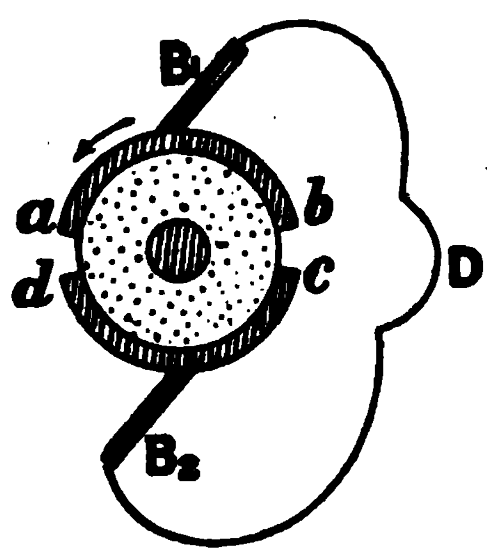


FIG. 504.

alternating current can most clearly be seen from Fig. 505, where A represents the manner in which the alternating current varies with the time, which is taken as abscissa, while at B the corresponding curve in the case of the rectified current is shown.

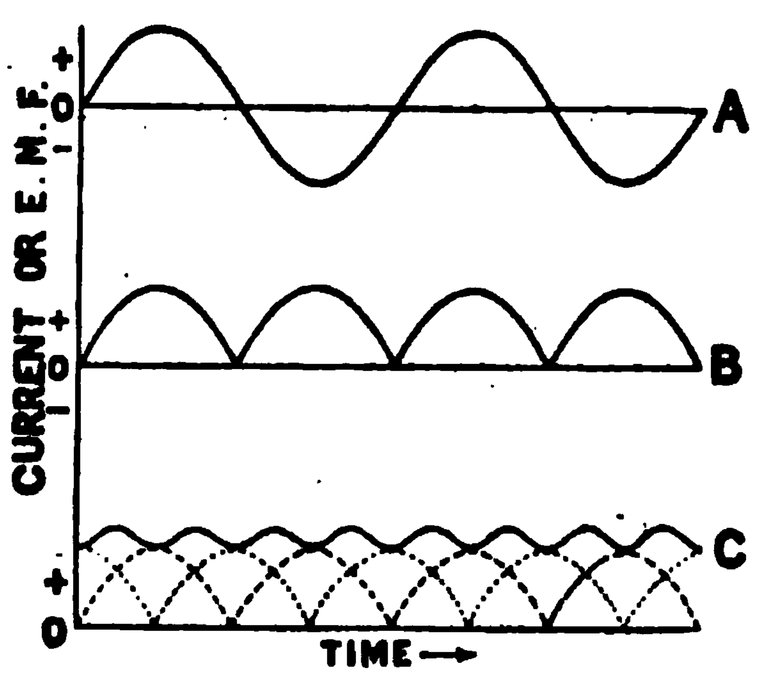


FIG. 505.

If a second coil of the same dimensions as the first were fixed to the same axle, so that its plane was at right angles to that of the first, and it were supplied with its own commutator, the brushes being connected to the same circuit as the first in such a way that the currents produced by the two coils in the external circuit were in the same direction, then the actual current in the circuit would be obtained by combining two such curves as that in Fig. 505 B. From the fact that one coil is placed a quarter of a revolution in advance of the other, the two curves must be displaced by a time equal to a quarter of a revolution, the one with respect to the other. In Fig. 505 C the dotted curves represent the currents due to the two coils separately, and the full-line curve the actual current due to the combined action of the two. It will be noticed how much more nearly uniform is the current than in the case where only one coil is used, and hence it will be understood how, by increasing the number of coils, what is practically a uniform current can be obtained.

527. Machines for the Conversion of Mechanical Energy into Electricity.—The arrangement described in the last section, although from its extreme simplicity it was useful as a means of explaining the production of the currents induced in a coil when rotated in a magnetic field, yet, on account of the weakness of any uniform field of the extent we there supposed and one produced in a space which was quite free from iron, the currents induced would only be very weak. In order to obtain stronger currents, it is necessary to have recourse to the use of iron in order to increase the induction through the rotating coil. Although the systematic description of even one or two of the different forms of machine which are used in practice for obtaining the strong currents which are now used is quite beyond the scope of this book, yet it may be of use to devote a few pages to considering the more general features which are more or less common to all.

In the first place, from a historical point of view rather than a practical one, such machines can be divided into two classes according to the means adopted for the production of the magnetic field in which the conductor in which the currents are induced is moved. Machines

in which the field is produced by means of permanent steel magnets are called magneto machines, while those in which the field is produced by electro-magnets are called dynamos.

The small machines which are used for the production of the currents of electricity used in medicine are examples of magneto machines. The field is produced by a horse-shoe magnet, while the coils in which the induced currents are generated are wound on soft iron cores. The coils and their cores are rotated near the poles of the magnet in such a way that the ends of the cores are brought alternately near the north pole and the south pole of the magnet. The result is that the cores become magnetised alternately in one direction and the opposite, and hence the induction through the coils which are wound over the cores is changed, being in one direction when the core is opposite the north pole, and in the opposite direction when the core is opposite the south pole. If required, the alternating currents thus produced are rectified by means of a commutator, such as was described in the last section.

528. Dynamo Electrical Machines.—In dynamo electrical machines the magnetic field is produced by means of electro-magnets which are magnetised by sending either the whole or part of the current produced by the machine round the coils of these magnets. The coil in

Fig. 506.

which the current is induced is called the armature, while the electro-magnets are called the field magnets. There are a great number of different forms of armatures in use, and we shall only describe the principles on which the action of three of these forms depend.

The Siemens armature consists of a coil of insulated wire wound longitudinally on a cylinder of soft iron as shown in Fig. 506. This armature is rotated between the poles of the field magnet NS, and as it rotates the induction through the coil changes in very much the same way as occurs in the simple coil considered in § 526, only the presence of the soft iron core on which the coil is wound very much increases the induction through the coil when it is placed in a given magnetic field. If a continuous current is required, a commutator is used to rectify the alternating current.

The Gramme armature is shown diagrammatically in Fig. 508, and the construction of an actual armature is shown in Fig. 507. This armature consists of a soft iron ring AA′ (Fig. 508) on which is wound a

continuous coil of wire. The commutator used consists of a number of copper bars, *m* (Fig. 507), which are separated from one another by some insulating material, usually mica. Each of these bars is connected with a point on the wire which is wound on the iron ring. The armature is capable of being rotated about an axis perpendicular to its plane between the poles, NS, of an electromagnet. On account of the greater permeability of the iron of the ring than that of the air or other non-magnetic materials between the poles, the lines of induction crowd through the iron in the manner shown in Fig. 509. Suppose now that the armature is rotated in the direction indicated by the arrow in Fig. 508, and consider one turn of the wire *abc* which is wound on the ring. In the position in which the turn *abc* is shown there are no tubes of induction pass through it. As, however, the armature rotates the number of tubes of induction passing through the coil increases till it reaches the position *def*. The result of the increase of the number of tubes of induction passing through the coil is to cause the production of an induced E.M.F. tending to send a current in the direction shown by the arrow. As the coil passes from *def* to *ghi* the number of tubes of induction which pass through it decreases, and an induced current in the reverse direction is produced. As the coil passes from *ghi* to *klm* the number of tubes which thread through it increases, but since they now pass through in the opposite direction, the induced E.M.F. is in the same direction as it was between *def* and *ghi*. Between *klm* and *abc* the number of tubes which pass through in this new direction decreases, and since this is the same thing as an increase in the number of tubes passing in the

FIG. 507.
(From Ganot's "Physics.")

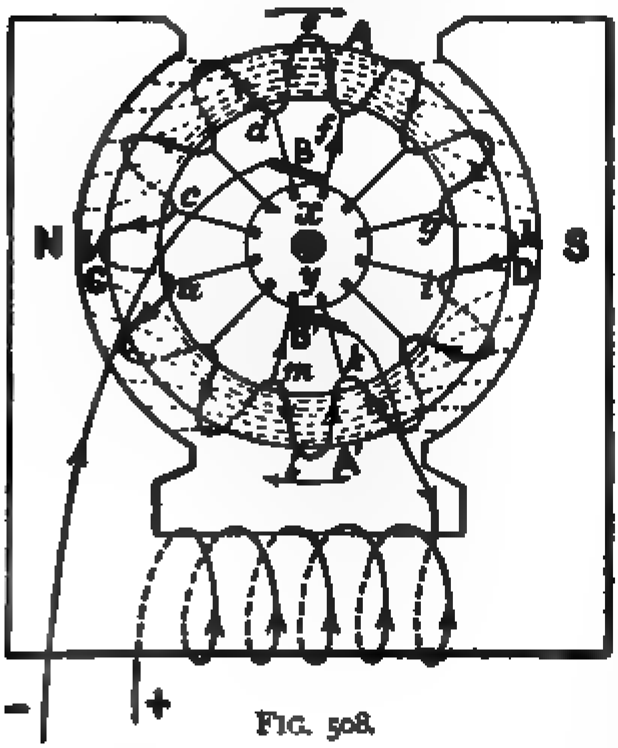


FIG. 508.

opposite direction, the induced E.M.F. will be in the same direction as it was while the ring was passing from *abc* to *def*. During the rotation of the armature each coil in succession goes through the same series of conditions as the one we have been considering, and the result is that the induced E.M.F. in the half ADA′ of the coils are all in the same direction, and so the actual induced E.M.F. between the points *f* and *m* is the sum of the E.M.F.'s induced in the separate coils between these points, while an equal and opposite E.M.F. is induced in the coils in the half A′CA. Since the induced E.M.F.'s in the two halves of the armature are equal and opposite, there is no E.M.F. tending to cause a current to circulate round the armature, although this consists of a closed circuit, but an E.M.F. is produced between the bars x and y of the commutator. Hence if two brushes, B, B′, make contact with the commutator at x and y respectively, and these brushes are connected to an external circuit, a current will be produced in this circuit. If the strength of the current produced by the machine is C, then each half of the armature is traversed by a current $C/2$.

The figure also shows the manner in which the current produced is used to magnetise the field magnets. When the machine is started, on account of the residual magnetism retained by the cast iron of which the cores of the field magnets are composed, there exists a weak field between the poles. The rotation of the armature in this field produces a small current, which traverses the coils of the field magnets and increases their magnetism, and this increase in the field increases the induced E.M.F., and hence also the current passing through the field magnets. This action of the induced current in increasing the strength of the field goes on, till, on account of saturation, the magnetisation of the magnets does not increase as the magnetising current increases.

In the description given above we have supposed that the lines of induction of the magnet which pass through the ring remain unaffected when the armature rotates and the machine produces a current. This, however, is not the case, for the current passing through the armature causes lines of induction to pass through the ring. Hence the form of the actual tubes of induction in the ring is obtained by compounding the field due to the field magnets with that due to the current in the armature. The result is that the tubes of induction have the form shown in Fig. 510, and as a result the points of the armature where the induction through the coil is a maximum, instead of being along the line AB, as we have supposed in our description, are along the line A′B′, being displaced in the direction in which the armature is rotating. The magnitude of this displacement of the points of maximum induction, and hence also of the positions where the brushes must touch the commutator, depends on the strength of the current the machine is sending, so that mechanism is usually provided to allow of the position of the brushes being adjusted.

Since in the Gramme armature the inside portion of each turn of the

wire on the armature moves in such a way, that, as shown in Fig. 508, it does not cut any lines of induction, or at any rate very few, this portion of the wire has very little beneficial effect as far as the production of an induced E.M.F. is concerned, while, since the induced current has to pass through this wire, electrical energy is wasted in heating the wire, according to

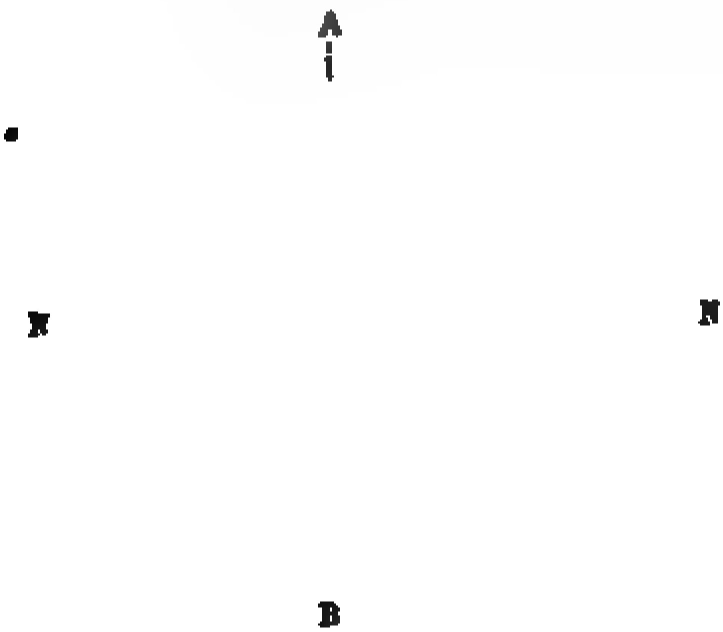


FIG. 509.
(From Ganot's "Physics.")

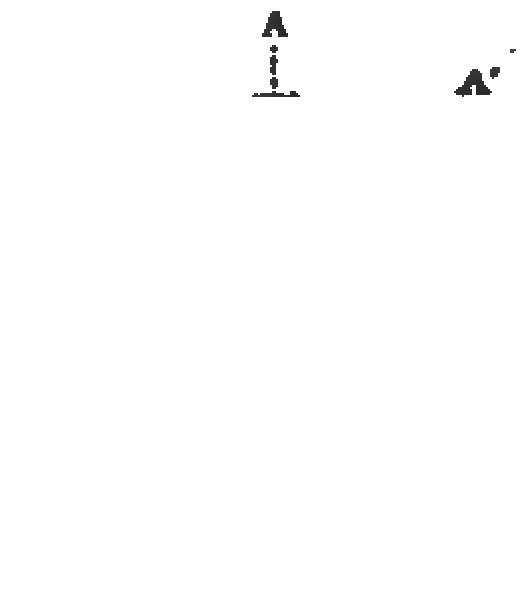


FIG. 510.
(From Ganot's "Physics.")

Joule's law. Hence the Gramme armature is better fitted for the production of small currents at a high potential than of very strong currents. On this account a different form of winding, called the drum winding, is adopted. In one form of armature a cylindrical core of soft iron is mounted so as to be capable of rotation about its axis between the poles of the field magnets, and is wound with wire in the manner shown diagrammatically in Fig. 511, in which, for simplicity, only four coils are shown, and each coil is supposed to consist of only one turn. In this arrangement, each turn, or set of turns, which builds up each section of the armature, is wound round the cylinder in very much the same way as the single coil in the Siemens armature. The ends of each turn or coil are brought to consecutive bars of the commutator, and the end of one turn or coil is connected to the same bar as the beginning of the next turn or coil. Taking any turn, it will be seen that the opposite sides are cutting through the tubes of induction in opposite directions, so that, as in the case of the simple coil considered in § 526, there is an E.M.F. acting round the coil, and both sides of the turn cut through tubes of

FIG. 511.

induction, and hence contribute to the induced E.M.F. The current is taken from the armature by means of two brushes making contact with the commutator.

In the case of each of the forms of dynamo considered, if a continuous current is sent through the field magnets and armature, this latter will be set in rotation, so that they will also function as electric motors for converting electrical energy into mechanical work. A consideration of the force which acts on a conductor conveying a current, when it is placed at right angles to the lines of force of a magnetic field, will at once show how it is these machines will act as motors.

529. Series, Shunt, and Compound Machines.—In the preceding section we have supposed that the armature and the coils of the field magnets were arranged in series, so that the whole of the current produced by the machine passes round the magnets. This arrangement is called a series machine. Since the E.M.F. induced in the armature is proportional to the strength of the magnetic field in which the armature turns, it follows that in a series machine the E.M.F., or voltage as it is called, increases as the current which the machine is furnishing increases, the speed of rotation being supposed constant. Since, as we have pointed out, the constancy of the voltage supplied to incandescent lamps is of great importance, this dependence of the voltage given by a series machine on the current being taken from it is an objection.

Another arrangement used is not to send the whole of the current which traverses the armature through the field magnets, but to let the field magnets form a shunt on the external circuit. This arrangement is called a shunt-wound machine, and in it part of the current supplied by the armature goes through the field magnets, and the rest through the external circuit. If now the resistance of the external circuit is reduced, so that the current sent by the machine increases, the proportion of the current which goes through the field magnet is reduced, for these are now shunted by a less resistance than before. Hence in this arrangement there is no tendency for the voltage to rise when the resistance of the external circuit is reduced, so that the machine is called upon to furnish a greater current. On the other hand, if the external resistance is very much reduced, the proportion of the current which traverses the field magnets is so small that the voltage will fall. For this reason a combination of the two kinds of winding is sometimes used, in which some of the magnetising current of the field magnets is supplied by a few turns of wire in which the whole current passes, these coils being in series with the armature, and the rest of the magnetising field is supplied by a number of turns of wire which are arranged in parallel with the external circuit. A machine wound on this principle is called a compound machine.

530. Back E.M.F. in Motors.—Suppose that the resistance of the armature and field magnets of a motor is R, and that it is connected to

a source of electromotive force, say a battery, which will produce a constant difference of potential of V volts at the terminals of the machine. Then if the armature is at rest, a current C, given by the equation $C = V/R$, will pass through the armature. If now the armature is set free, so that it is allowed to revolve, then, since if the armature were driven round in the same direction as that in which it turns an E.M.F. would be developed at its terminals in the opposite direction to that which is used to drive it, it follows that the armature by its motion will create an induced E.M.F. which will oppose the E.M.F. V which is sending a current through the motor. This counter E.M.F., as it is called, will increase as the speed of the motor increases, since the induced E.M.F. depends on the speed with which the conductors on the armature cut through the tubes of induction of the field. Let v be the counter E.M.F. developed at any given speed, then the effective E.M.F. sending a current through the machine is $V-v$, and hence the current which traverses the armature is given by

$$C = (V - v)/R.$$

If the machine is supposed to turn without friction and to do no external work, the speed will go on increasing till the counter E.M.F. is equal to V. Under these conditions there will now be no force acting on the armature tending to make it rotate, and hence, since we have postulated the absence of friction, the machine will continue to turn at a constant speed. If now the machine is caused to do external work, say to wind up a weight, then the speed will decrease, and the back E.M.F. will decrease, so that a current will pass through the machine.

Suppose that the power developed by the machine, that is, the rate at which it does work, is P, and that either the friction in the different parts of the machine is so small as to be negligable, or, what comes to the same thing, that the power P includes the work done against friction. If then the back E.M.F. is v, the current passing through the machine will be $(V-v)/R$. The energy corresponding to this current will be spent partly in heating the wire forming the armature, and partly in doing the work P. The part of the energy spent in producing heat is, by Joule's law, C^2R or $(V-v)^2/R$. Since the E.M.F. between the terminals of the machine is V, the energy supplied by the current C in one second when flowing through this drop of potential is CV. Hence the energy available for doing external work is

$$\frac{(V-v)V}{R} - \frac{(V-v)^2}{R},$$

or

$$P = v(V-v)/R.$$

From this expression it will be seen that P is zero, that is, the machine does no external work, both when $V = v$ and when $v = 0$. The

first case is that which we have already considered, when the motor revolves at such a speed that the back E.M.F. is equal to the applied E.M.F. The other case, when $v=0$, is when the armature is at rest, and when the current is V/R, and hence the heat developed according to Joule's law is C^2R or VC, that is, is equal to the energy supplied by the external source, so that there is none available for doing external work. If the speed of the motor is by some external means increased, so that v is greater than the applied E.M.F., the motor will operate as a generator and will send a current in the reverse direction round the circuit, and in this way will supply energy to the circuit.

The power developed by the motor will be a maximum when $V=2v$. Then the power given by the motor is V^2/R, and the energy supplied will be $2V^2/R$; so that the power developed will be a maximum, when the speed is such that the back E.M.F. is half the applied E.M.F., and half the energy supplied will be converted into useful work and half wasted in heat. It must, however, be noted that although this is the speed for which, having given the external E.M.F., most work can be done by the motor, it is by no means the most economical speed at which to run the motor. The energy supplied is $V(V-v)/R$, while the energy converted into work is $v(V-v)/R$. Hence the ratio of the energy converted into useful work to the energy supplied is $v(V-v)/V(V-v)$ or v/V. Thus the proportion of the energy supplied which is converted into useful work increases as v is made more nearly equal to V. As we have seen, however, as the speed is increased, so that v may become more nearly equal to V, although the proportion of the energy supplied which is converted into useful work is large, yet, since the amount of energy which the motor is then capable of taking from the external circuit is very small, the power developed must also be small. In practice it is usual to run motors at speeds so much above that for which $V=2v$, that nearly 90 per cent. of the energy supplied is converted into useful work.

531. Alternating Currents—Transformers.—The employment of electricity for the transmission of the energy developed, say, at a waterfall, to the neighbouring towns, over distances of many miles, makes the question of the cost of the conductors employed to convey the current from the generating point to the place where it is used of considerable importance. Suppose that it is required to transmit power from A to B, so that the energy available at B is W watts. If R is the resistance of the conductor extending from A to B and back, by means of which the current is conveyed, and C is the current transmitted, while V is the E.M.F. between the wires at the generating end. Then by Ohm's law the fall of potential along the wires will be equal to RC, and hence the E.M.F. available at B will be $V-RC$. Thus the energy available at B will be $C(V-RC)$, and this is to be equal to W. The watts wasted in heat in the conducting wires is by Joule's law C^2R.

Hence the object is to make C^2R as small as possible, while keeping $C(V-RC)$ constant. One way of doing this is to reduce the value of R, that is, to increase the diameter of the wire used to convey the current. This, however, involves a great outlay on copper. Another way of reducing the loss of energy in the conducting wires is to reduce the current C, but under these circumstances, if W is to remain constant, V must be made large; that is, a great potential difference between the wires must be employed. Since, however, an accidental contact with the wires conveying currents at high potentials is fatal to life, such currents are not suitable for use in houses for lighting purposes, or for driving machinery in workshops. There is a further difficulty, that to produce directly such high potential currents involves very complete insulation between the separate turns of the armature of the dynamo employed. It is thus evident that if by any means the low-tension current produced by the generator were converted into a high-tension current, and this current were transmitted to the distant station where it was again converted into a low-tension current, the advantage of the small loss of energy during the transmission with the absence of the danger attached to the use of high-tension currents would be attained. In the case of continuous currents, this transformation from low to high tension and *vice versa* is not possible, except by virtually using a motor driven by the one current to drive a dynamo to produce the other, but with alternating currents the case is quite different.

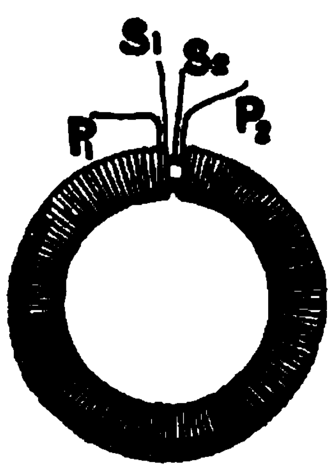


FIG. 512.

Suppose that an iron ring (Fig. 512) is lapped over with a layer of insulated wire, there being N turns, the cross-section of the iron being s and the axial length of the ring l. Then if a current C is sent through this coil, we shall have an induction B in the iron equal to $4\pi N\mu C/l$. For suppose that the iron were removed, then by § 475 the work which would be done in carrying a unit pole once round the inside of the annular space which was occupied by the iron will be $4\pi NC$, for it will have been carried once round N conductors, in each of which a current C is flowing. If H is the strength of the field within the coil produced by the current, then the work done in carrying a unit pole once round will be lH. Hence, equating these two values, we get the value of the magnetising field when the iron is placed within the coil as

$$H=4\pi NC/l.$$

But if μ is the permeability of the iron for the magnetising field H, then the induction B is given by $B=\mu H$, or, substituting for H the value just obtained,

$$B=\frac{4\pi\mu NC}{l}.$$

The total number of tubes of induction which pass through the cross-

section of the iron is sB, so that the number of tubes of induction which pass round the ring of iron is given by

$$sB = \frac{4\pi\mu sNC}{l}.$$

If in addition a second coil containing n turns of wire is lapped round the ring, then each turn of this coil will be traversed by sB tubes of induction, or the whole coil will be traversed by nsB tubes. If now, instead of the current being constant it is an alternating current, and if C is the maximum value of the current, the induction through the secondary coil will vary from $+nsB$ to $-nsB$, and an induced E.M.F. will be produced. If R is the resistance of the primary coil, then by making R small, that is, having a few turns of thick wire, the applied E.M.F. required to send the current C through the primary coil may be made small. On the other hand, since the induction sB takes place through each of the turns of the secondary, and that the induced E.M.F. in the whole coil is the sum of the induced E.M.F.'s in the separate turns, by making the number of turns, n, in the secondary large the induced E.M.F. may be made large. Hence by sending an alternating current through one of the coils an alternating induced current will be produced in the other coil, and by suitably varying the ratio of the number of turns in the two coils the induced E.M.F. may be made to bear any required relation to the E.M.F. used to send the current in the primary circuit. The exact relation between the primary and the secondary E.M.F.'s is complicated by the effects of self and mutual induction as well as by the hysteresis of the iron. The above will, however, explain the general principles on which transformers, as such arrangements are called, work. It will be noticed that an induction coil is simply a transformer in which the secondary has a relatively great number of turns, so that the E.M.F. induced in it is very great.

In the employment of transformers for the transmission of power the generating dynamo gives a relatively low voltage, and by means of a transformer the current produced is transformed into a high-pressure current, which is transmitted to the place where the electrical energy is to be used. Here, by means of a second transformer, the current is again converted into a low-pressure current.

532*. The Magnetic Circuit.—In the last section we showed that if we have a soft iron annulus of cross-section s, which is lapped round uniformly with N turns of wire, the length of the iron core measured along the axis being l, the total induction through the iron is given by

$$Bs = 4\pi NC\mu s/l,$$

where C is the current flowing in the wire. If we call the total induction through any cross-section of the iron G, we may write—

$$G = 4\pi NC.\mu.s/l.$$

Now if the iron were removed, the work which would have to be done to carry the unit pole once round the annular space which was occupied by the iron would be $4\pi NC$, for we should have carried the unit pole once round N circuits, in each of which a current C was flowing. Hence, if we call this quantity of work M, we have

$$G = M.\mu.s/l.$$

Thus the total induction through the iron is obtained by multiplying M by a factor, μ, which depends on the nature of the material (iron) forming the annulus, and by a factor, s/l, which depends on the geometrical dimensions of the portion of the substance considered. Now if an E.M.F., E, acts between the ends of a uniform conductor, of which the cross-section is s and the length l, and of which the specific conductivity is k, the current C flowing in the conductor is given by

$$C = E.k.s/l.$$

The E.M.F., E, acting between the ends of the conductor is measured by the work which is done on unit quantity of electricity as it flows from one end of the conductor to the other, so that the current is equal to the product of this quantity of work into a factor, k, which depends on the nature of the conductor, and into a factor, s/l, depending on the geometrical dimensions of the conductor. Stated in this way there will be seen to be a certain parallelism between the magnetic equation and the electrical equation, and this parallelism has led to the adoption in the magnetic case of a terminology suggested by the electrical problem. Thus the quantity M, which plays the same part in the magnetic equation as does the E.M.F. in the electrical problem, is called the *magneto-motive force*, and μ has been called the specific magnetic conductivity. In the same way, since the factor l/sk represents the resistance of the conductor, the factor $l/s\mu$ has been called the magnetic resistance or reluctance. Using this terminology, we have that the total magnetic induction through a magnetic circuit[1] is equal to the magneto-motive force divided by the reluctance.

Although this manner of viewing magnetic problems is of considerable use, particularly when dealing with practical problems, such as the design of dynamos and transformers, and has proved suggestive in indicating new paths for experimental research, it must be remembered that the whole analogy is a mathematical one, built up on the similarity of the two equations considered above, there being no physical analogy between the two cases. Thus there is no known magnetic phenomenon which is

[1] We have already seen that every tube of induction is an endless tube; thus the portion of space through which any tube, or set of tubes, passes in their whole length forms a closed circuit, and it is therefore known as a magnetic circuit. A magnetic circuit may be formed by one or more different media, and may be single or branched, just as an electrical circuit may be formed by different substances, and may have branches forming loops.

physically analogous to the conduction current, while physically the analogue of permeability is not specific conductivity but specific inductive capacity. Again, while the resistance of a conductor is independent of the strength of the current, the reluctance depends on the magnetic induction, for, as we have seen in § 504, the permeability of iron varies enormously with the induction.

In the case of the annulus considered above, the tubes of induction are confined to the iron, and the magnetic circuit therefore consists of one medium only. We may, however, apply the idea of the magnetic circuit to cases where the tubes of induction pass through media of different permeability.

In the first place, let us take the case of the iron ring already considered, but suppose that the magnetising coil, instead of being wound uniformly all round the ring, is confined to a small portion of the circumference. Under these circumstances, some of the tubes of induction will leave the iron in the part of the ring which is not covered by the magnetising coil, and will travel through the air. Since, however, the permeability of soft iron is several hundred times greater than that of the air, at any rate when the magnetising field is not very great, such a large proportion of the tubes of induction will continue all the way through the iron ring, that we may, without making any appreciable error, neglect the ones that do not. If the iron were removed, and a unit pole were carried once round the space previously occupied by the iron ring, it would pass once round each of the turns of the magnetising coil. Hence if there are N turns in this coil, and the current is C, the work done is $4\pi NC$, and therefore the magneto-motive force is $4\pi NC$. Also the reluctance of the iron ring is $l/s\mu$. Thus the total induction through the iron is given by

$$G = \frac{\textit{Magneto-motive force}}{\textit{Reluctance}} = 4\pi NCs\mu/l.$$

This result is slightly greater than that in the iron which is furthest from the magnetising coil, on account of the tubes of induction which thread through the coil, but instead of passing through the iron, pass through the surrounding air. Still the result obtained is a very near approximation to the truth. The advantage of the magnetic-circuit point of view is apparent if we consider how very difficult it would be to calculate the value of the magnetising force at each point of the iron ring, in order to deduce the induction. The analogue of this problem in electricity would be the case of a ring of copper immersed in a feebly conducting medium, such as water, for in such a case most of the current would traverse the copper, but some would traverse the water, and so the resistance of the circuit would be somewhat less than the resistance of the copper alone, although a very near approximation to the current would be obtained if we neglected the portion of the current which flows through the water.

As another example of the utility of the idea of the magnetic circuit, we may take the case of the iron ring which is lapped over with a uniformly wound magnetising coil, but which at one place has been cut so that the continuity of the iron is broken by a narrow air-gap. As before, the magneto-motive force will be $4\pi NC$. The magnetic circuit is no longer confined to a single medium, but at the gap passes from iron to air. Hence in calculating the reluctance, we have to consider the two portions of the circuit, in one of which the permeability is μ, and in the other portion it is unity. If x is the width of the gap, the length of the iron circuit is $l-x$. Hence the reluctance of the iron part of the circuit is $(l-x)/\mu s$. If the air-gap is at all wide, the tubes of induction will spread out at the gap, and hence the cross-section of the magnetic circuit in the gap will be greater than s. If, however, the gap is very narrow, the spreading of the tubes will be inappreciable, and we may take the cross-section of the circuit at the gap as equal to s. The length of the air part of the circuit being x, the reluctance is x/s. Hence the reluctance of the combined iron and air circuit is

$$(l-x)/\mu s+x/s.$$

Thus the total induction through the circuit is given by

$$\begin{aligned} G&=4\pi NC/\{(l-x)/\mu s+x/s\} \\ &=4\pi NC/\left\{\frac{l+x(\mu-1)}{\mu s}\right\} \\ &=4\pi NC\mu s/\{l+x(\mu-1)\}. \end{aligned}$$

If the length of the ring had been $l+x(\mu-1)$, and supposing no gap were present, the induction would have been

$$4\pi NCs/\{l+x(\mu-1)\}.$$

Hence the effect of a gap of length x in reducing the induction is the same as would be produced by a length $x(\mu-1)$ of iron. If the magnetising field ($4\pi NC$) is 5 *c.g.s.* units, the permeability for the soft iron, for which the curve in Fig. 481 is drawn, is 2400. Hence if the length of the ring is 30 cm., and its cross-section 4 sq. cm., the induction with a gap a millimetre wide is

$$\frac{5\times2400\times4}{30+0.1(2400-1)}=\frac{48000}{269.9}=177.8 \text{ unit tubes.}$$

If no gap were present the induction would be

$$\frac{5\times2400\times4}{30}=1600 \text{ unit tubes.}$$

It will thus be seen how enormously the presence of the air-gap reduces the total induction.

533. The Electric Telegraph.—Since the direction in which a galvanometer needle is deflected depends on the direction of the current which is sent through it, by having an arrangement at one station by which the current sent by an electric battery can be reversed in direction, and connecting the commutator with a galvanometer placed at the other station by an insulated conducting wire, signals can be transmitted from the battery station to the other. Only one conducting wire is in general used, the earth being used for completing the circuit. By having a battery and a galvanometer at each station, which by means of keys can be connected to the circuit and the current reversed, messages can be sent in both directions. The older forms of electric telegraph were on this principle, the receiving instruments being, in fact, somewhat unsensitive galvanometers in which the deflection of the needle to right and left was observed. At the present time nearly all telegraphy is done by means of the Morse sounder. This instrument consists of a small electro-magnet, through the coils of which the current sent from the sending station is passed. This current causes the electro-magnet to attract a light soft iron armature, which is held away from the pole of the electro-magnet by means of a spring. The armature, when it strikes the pole, makes a distinct click, and from the number of clicks and the interval between them, the operator reads the signal. The current is sent by means of a key, on the depression of which the circuit of the battery is completed. In the following table the code ordinarily employed, and called the Morse code, is given. A long stroke means that the interval between that click and the next has to be longer than that between a short stroke and the next.

International

THE ~~MORSE~~ ALPHABET.

A	- ——	J	- —— —— ——	S	- - -
B	—— - - -	K	—— - ——	T	——
C	—— - —— -	L	- —— - -	U	- - ——
D	—— - -	M	—— - - -	V	- - - ——
E	-	N	—— -	W	- —— ——
F	- - —— -	O	—— —— ——	X	—— - - ——
G	—— —— -	P	- —— —— -	Y	—— - —— ——
H	- - - -	Q	—— —— - ——	Z	—— —— - -
I	- -	R	- —— -		

When the distance between the sending and the receiving stations is considerable, on account of the resistance of the connecting wire, the current which can be sent is not sufficiently strong to attract the armature of the receiving instrument and make it give an audible sound. In these circumstances what is called a relay is employed. This consists of an electro-magnet, round the coils of which the current which

is sent from the distant station is sent. When the current passes through this electro-magnet it attracts a very light and delicately poised armature. This armature works between two stops, and when it is attracted by the magnet against the one stop, it completes the circuit of a local battery which has the sounder in its circuit. Hence the sounder is worked by a battery at the station at which the signal is received, and the current transmitted from the distant station is only used to complete the circuit of the local battery.

In the case of submarine telegraphy, where the distances between the stations are often very great, the receiving instrument is practically a very sensitive mirror galvanometer, and the message signals are formed by deflections of the spot of light reflected from the mirror to right and left, a deflection to the right corresponding to a dash in the Morse alphabet, and a deflection to the left to a dot.

In duplex telegraphy two messages are sent simultaneously through the same wire, one in each direction. This is accomplished by winding the receiving instrument G (Fig. 513) with two coils, which are so arranged that when the battery B is connected to the circuit by pressing the key K so as to rest on the stud *a*, the current which passes from the battery divides at the instrument, part going through one coil and part through the other, and in such a direction that the effects of the currents in the two coils on the needle of the instrument are in opposite directions. One coil of the instrument is connected to the line which goes to the other station, while the second coil is connected through a variable resistance R with the other pole of the battery and the plate E, which is buried in the earth. If then the resistance of the one coil, the line, the receiving instrument at the other station, and of the return circuit through the earth, is the same as that of the second coil of the instrument at the sending station, together with the resistance R, the current which passes from B will divide into two equal parts; and since these parts traverse the coils of G in opposite directions, they will exactly neutralise each other's effect on the instrument G, so that the working of the key K will not affect the instrument G. The current sent through the line will, however, only traverse one of the coils of the instrument at the other station, and hence it will affect this instrument.

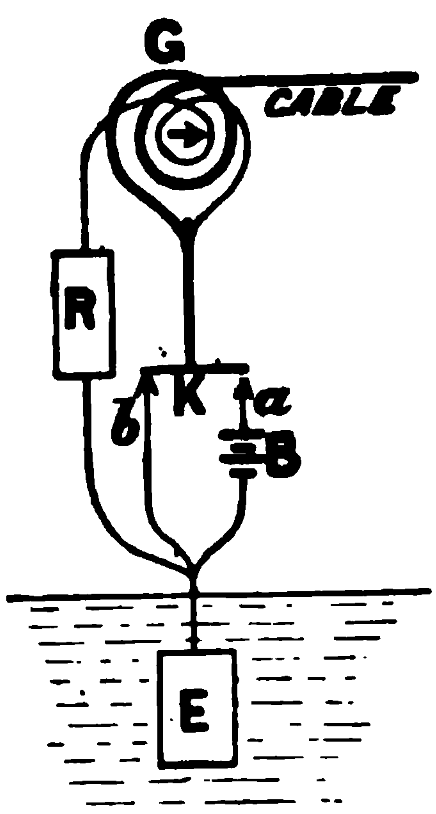


FIG. 513.

There are other systems of duplex telegraphy, as well as methods by means of which more than two simultaneous messages may be sent through the same line, but space will not allow of these being considered.

534. The Telephone.—The telephone was invented by Graham Bell, and a section of a Bell telephone receiver is shown in Fig. 514. It consists of a steel bar-magnet, M, fitted inside a case, L. A coil, B, of a large number of turns of fine wire fits over one pole of the magnet, the ends of the coil being connected with the terminals C. At a distance of about a millimetre from the pole of the magnet on which the coil is wound is fixed a diaphragm, D, composed of a sheet of thin soft iron. This diaphragm is held in position by being clamped between the mouth-piece E and the case L.

The diaphragm becomes magnetised by induction, and this induced magnetisation reacts on the permanent magnetisation of the magnet M, the amount of the reaction being dependent on the distance of the centre of the diaphragm from the surface of the pole. When the instrument is spoken into, the vibrations of the air cause the diaphragm to vibrate in unison, and by its to-and-fro motion the diaphragm causes the

FIG. 514.
(*From Ganot's "Physics."*)

magnetisation of the magnet to vary also in unison with the incident air-vibrations. The changes of the strength of the magnet mean that the number of tubes of induction passing through the coil B must also vary, and hence a series of induced currents are produced in a circuit of which this coil forms a part. If the terminals C are connected by wires to a second instrument, the induced currents which are produced by the motion of the diaphragm D will traverse the coil of the second instrument, and will produce a change in the magnetisation of the magnet in this instrument. These changes in the strength of the magnet in the second instrument will cause changes in the force with which the magnet attracts its diaphragm, and hence this diaphragm will be set in vibration in such a way as to reproduce the vibrations which were produced in the diaphragm of the first instrument; and in this way the air in the neighbourhood of the diaphragm will be set in vibration, and the sounds produced near the transmitting instrument will be reproduced.

The amplitude of the excursions of the telephone diaphragm are

excessively small. Thus Barus has measured it and found it to be about 10^{-6} cm. when the instrument is emitting a sound which is just audible. The currents which are produced are also very small, being about 2×10^{-4} ampere in the case of the ordinary transmission of speech.

535. The Microphone.—The microphone, an instrument invented by Professor Hughes, consists essentially of an arrangement by which one part of a circuit, in which is included a telephone and an electric battery, is completed by two conductors which rest lightly the one on the other. One form of the microphone is shown in Fig. 515. A piece of gas carbon, D, pointed at each end, rests lightly in two small cup-shaped hollows made in two pieces of the same kind of carbon, C, C′. The rod D is not clamped between the other rods, but rests on the lower one, and is prevented from falling by the upper end resting against the side of the cup made in the upper rod. The terminals of the circuit containing the battery and telephone are attached to the rods C, C′, by the wires A and B. When a disturbance is produced, such as by the ticking of a watch placed on the base of the instrument, the rod D is set in motion, so that the pressure with which it rests against the upper rod varies. Since the resistance of carbon changes with the pressure to which it is subjected, the movements of D cause variations in the resistance of the circuit in which the microphone is included, and these changes in resistance cause corresponding changes in the current which traverses the circuit, these variations in the current causing the telephone to sound. The sensitiveness of the instrument is very great, so that even a very minute disturbance produced near the base of the instrument, such as the noise made by a fly walking on the wood, is enough to cause the telephone to reproduce the noise in quite an audible form.

Fig. 515.

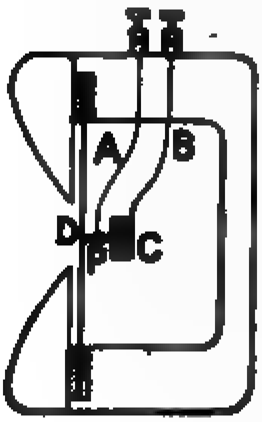


Fig. 516.

The principle of the microphone has been applied to replace the telephone as a means of producing the changing currents required to transmit speech, an ordinary telephone being used to reproduce the sounds. The Blake form of microphone transmitter consists of a sheet-iron diaphragm, D (Fig. 516), held in position behind a mouthpiece by being clipped between two rubber bands. A small piece of platinum wire, P, is attached to a slender spring, and bears at one end on the centre of the diaphragm, and at the other end against a piece of gas carbon, C, which is itself carried by a spring, B. The springs A and B are connected to a circuit in which are included a battery and the telephone at the receiving station. When the mouthpiece is spoken into the diaphragm is set in vibration, and so causes the platinum

P to press on the carbon block with a variable pressure, and in this way the resistance, and hence also the current which traverses the instrument, varies in unison with the motion of the diaphragm. In this form of transmitter the energy necessary to produce the motion of the diaphragm at the distant station is supplied by the battery, and is not, as is the case when a telephone is used as transmitter, derived from the energy of motion of the receiving diaphragm. Hence in the carbon transmitter the receiving diaphragm only has to control the supply of energy of the battery, and so plays the part of the relay used in long-distance telegraphy.

536. Dimensions of Electrical and Magnetic Quantities.—In the preceding pages nothing has been said as to the dimensions of the different electrical and magnetic quantities with which we have been dealing. Two systems of units have been employed, the one, the electro-static system, in which the fundamental quantity was the quantity of electricity, and was defined by means of the repulsion exerted between two charges when in air; and the other, the electro-magnetic system, in which the fundamental quantity was the unit pole. Taking first the case of the electro-static system of units. If two charges of Q_s[1] units are placed at a distance d apart in air, the force F with which they act the one on the other is, by the definition which we have adopted for Q, given by the equation $F=Q_s^2/d^2$. If, however, instead of being placed in air the charged bodies are placed in a medium of which the specific inductive capacity is K, the force exerted between the charged bodies is given by $F=Q_s^2/Kd^2$. Now, just as in the case of temperature considered in § 265 we were not able to determine the dimensions of temperature in terms of the fundamental units of length, mass, and time, and hence were obliged to keep in our dimensional equations a symbol to represent the unknown dimensions of temperature, so in the electrical case we do not know the physical nature of specific inductive capacity, and cannot therefore determine its dimensions in terms of the fundamental units, and have to indicate in the dimensional formulæ the unknown dimensions of specific inductive capacity by a symbol, K. The reason the quantity K does not come into the ordinary expression given for the definition of the unit quantity of electricity on the electro-static system, is that we make the perfectly arbitrary assumption that the specific inductive capacity of air is unity. Of course, if the specific inductive capacity were independent of the units of mass, length, and time, and not simply apparently independent on account of our want of knowledge of the true nature of specific inductive capacity, then the dimensions of K would be zero, and would not appear in the dimensional equations. Many writers on this subject assume, although there is no

[1] We shall use a subscript s to indicate that the quantity is measured in electro-static units. Similarly a subscript m will be used to indicate that the quantity is measured in electro-magnetic units.

particle of justification for such an assumption, that the dimensions of K are zero, and then proceed to develop the dimensional equations for the different electrical quantities. The result of such an assumption is that we are led to equations between different electrical and magnetic quantities, the two sides of which appear as if they were of different dimensions, a result which is on the face of it absurd.

Taking the symbol K to represent the unknown dimensions of specific inductive capacity, since $Q_s^2 = d^2FK$, we have the dimensional equation $[Q]=[LF^{\frac{1}{2}}K^{\frac{1}{2}}]$. Since F is a force, $[F]=[LMT^{-2}]$. Hence

$$[Q_s]=[L^{\frac{3}{2}}M^{\frac{1}{2}}T^{-1}K^{\frac{1}{2}}].$$

Since the potential, V_s, at a point is equal to the work which has to be done in moving the unit charge from a point where the potential is zero to the given point, or is the work W done in moving a charge Q_s divided by Q_s, we have the equation

$$V_s = W/Q.$$

Hence the dimensional equation is

$$[V_s]\,[W^1Q_s^{-1}],$$

or, substituting the dimensions of Q_s and of W $([L^2M^1T^{-2}])$, we have

$$V_s=[L^2M^1T^{-2}]/[L^{\frac{3}{2}}M^{\frac{1}{2}}T^{-1}K^{\frac{1}{2}}]=[L^{\frac{1}{2}}M^{\frac{1}{2}}T^{-1}K^{-\frac{1}{2}}].$$

The capacity, M, of a conductor being equal to the charge divided by the potential which this charge produces, or

$$M=Q/V,$$

we get the dimensional equation

$$[M_s]=[Q_sV_s^{-1}],$$

or $$[M_s]=[L^{\frac{3}{2}}M^{\frac{1}{2}}T^{-1}K^{\frac{1}{2}}]/[L^{\frac{1}{2}}M^{\frac{1}{2}}T^{-1}K^{-\frac{1}{2}}]=[LK].$$

Thus the capacity of a conductor depends on the units of length and of specific inductive capacity only. This agrees with the result which we obtained in § 464, for we there found that the capacity of a sphere at a great distance from all other conductors is, in air, numerically equal to the radius, which of course only involves the unit of length.

The current, C_s, passing along a conductor is measured in the electrostatic system by the quantity of electricity, Q_s, which passes through the conductor in the unit of time, hence

$$C_s=Q_s/T,$$

or $$[C_s]=[QT^{-1}]=[L^{\frac{3}{2}}M^{\frac{1}{2}}T^{-2}K^{\frac{1}{2}}].$$

In the electro-magnetic system the fundamental quantity is the unit magnetic pole, which is defined in such a way that if two poles, each of

strength m, are placed at a distance d apart in air, the force, F, which they will exert one on the other is given by the equation $F=m^2/a^2$. If, instead of being in air, the poles are separated by a medium of which the permeability is μ, the force exerted between them is given by

$$F=m^2/\mu d^2.$$

In this case we are unable to express the permeability μ in terms of the fundamental units, and hence in the dimensional equations we have to keep in a symbol to represent the dimensions of μ. Hence if we indicate the unknown dimensions of μ by $[\mu]$ we get the dimensional equation

$$[m]=[LF^{\frac{1}{2}}\mu^{\frac{1}{2}}]=[L^{\frac{3}{2}}M^{\frac{1}{2}}T^{-1}\mu^{\frac{1}{2}}].$$

The magnetic force H_m at a point, or the strength of the magnetic field at the point, is the force which acts on the unit pole when placed at the point. Hence if F is the force acting on a pole of strength m when placed at the point, we have

$$F=mH_m.$$

Thus $$[H_m]=[Fm^{-1}]=[L^{-\frac{1}{2}}M^{\frac{1}{2}}T^{-1}\mu^{-\frac{1}{2}}].$$

Since the magnetic induction, B, is connected with the strength of the magnetising field by the relation $B=\mu H$, the dimensions of B are given by

$$[B_m]=[H_m\mu]=[L^{-\frac{1}{2}}M^{\frac{1}{2}}T^{-1}\mu^{\frac{1}{2}}].$$

The magnetic moment, M', of a magnet being the product of the strength of the pole into the length, l, the dimensions are given by

$$[M'_m]=[mL]=[L^{\frac{3}{2}}M^{\frac{1}{2}}T^{-1}\mu^{\frac{1}{2}}]\times[L]=L^{\frac{5}{2}}M^{\frac{1}{2}}T^{-1}\mu^{\frac{1}{2}}].$$

Intensity of magnetisation, I, being the magnetic moment per unit of volume, we have

$$[I_m]=[M'_m]/[L^3]=[L^{\frac{5}{2}}M^{\frac{1}{2}}T^{-1}\mu^{\frac{1}{2}}]/[L^3]=[L^{-\frac{1}{2}}M^{\frac{1}{2}}T^{-1}\mu^{\frac{1}{2}}].$$

The susceptibility, k, is equal to the quotient of the intensity of magnetisation by the magnetising force. Hence

$$[k]=[I_m]/[H_m]=[L^{-\frac{1}{2}}M^{\frac{1}{2}}T^{-1}\mu^{\frac{1}{2}}]/[L^{-\frac{1}{2}}M^{\frac{1}{2}}T^{-1}\mu^{-\frac{1}{2}}]=[\mu].$$

The force, F, exerted upon a pole of strength m when placed at the centre of a circle of radius r, when a current C flows along a length l of the circumference of this circle, is given by

$$F=mCl/r^2.$$

Hence the dimensions of current are given by

$$[C_m]=[FLm^{-1}]=[L^2MT^{-2}m^{-1}]=[L^{\frac{1}{2}}M^{\frac{1}{2}}T^{-1}\mu^{-\frac{1}{2}}].$$

If E is the electromotive force between two points on a conductor,

then the power, P, which has to be used to cause a current C to flow against this E.M.F., is given by

$$P=EC.$$

Hence

$$[E_m]=[P]/[C_m]=[L^2MT^{-3}]/[L^{\frac{1}{2}}M^{\frac{1}{2}}T^{-1}\mu^{-\frac{1}{2}}]=[L^{\frac{3}{2}}M^{\frac{1}{2}}T^{-2}\mu^{\frac{1}{2}}].$$

The resistance of a conductor is the ratio of the difference of potential between the ends, when traversed by a current C, to that current. Hence

$$[R_m]=[E_m]/[C_m]=[L^{\frac{3}{2}}M^{\frac{1}{2}}T^{-2}\mu^{\frac{1}{2}}]/[L^{\frac{1}{2}}M^{\frac{1}{2}}T^{-1}\mu^{-\frac{1}{2}}]=[LT^{-1}\mu].$$

Since quantity of electricity is equal to the product of the current passing into the time, we have on the electro-magnetic system

$$[Q_m]=[C_mT]=[L^{\frac{1}{2}}M^{\frac{1}{2}}\mu^{-\frac{1}{2}}].$$

Proceeding in this way, we may find the dimensions of any of the other electrical or magnetic quantities on either the electro-static or electro-magnetic system.

We have seen in § 502 that the induction, magnetising force, and the intensity of magnetisation are connected together by the equation $B=H+4\pi I$. Now, since each term of any physical equation must be of the same dimensions as the other terms, this equation indicates that B, H, and I are all of the same dimensions. But by definition $B=\mu H$, so that the dimensions of B and H must be different if μ is not a simple number having no dimensions. The reason for this apparent anomaly is that the equation between B, H, and I given above is not a general equation, but only holds when the magnetic body is surrounded by a medium, such as air, of which the permeability is taken arbitrarily as unity. The permeability of air is taken as unity simply because our knowledge as to the nature of permeability is not sufficient to tell how to measure it in a way independent of the properties of any one kind of matter.

If, instead of taking the permeability of air as unity, we call it μ_0, that is, we use the symbol μ_0 to indicate the permeability measured in absolute units, although, on account of our ignorance of the true nature of magnetism, we are unable at present to say how it is to be measured, then the equation connecting B, H, and I can be shown to be

$$B/\mu_0=H+4\pi I/\mu_0.$$

In this equation B and I refer to a medium of absolute permeability. μ and μ_0 is the absolute permeability of the medium in which H is measured. In this equation all the terms are of the same dimensions, so that while the dimensions of B and I are the same, the dimensions of B and H differ by the dimensions of μ. If in this equation we assume arbitrarily that the permeability of the medium (air) in which

H is measured is unity, we get $B=H+4\pi I$. When treating of dimensions, however, it is not allowable to assume that the *dimensions* of permeability is zero, so that the corresponding dimensional equation must always include a symbol to represent the unknown dimensions of the absolute permeability of air.

537. Connection between the Two Sets of Units.—In the preceding section we have obtained the dimensions of quantity of electricity in both the electro-static and the electro-magnetic systems. Since the dimensions of any physical quantity must be independent of the particular system of units adopted, we may equate the two values for the dimensions of electrical quantity, and we thus obtain the following equation :—

$$[L^{\frac{3}{2}}M^{\frac{1}{2}}T^{-1}K^{\frac{1}{2}}]=[L^{\frac{1}{2}}M^{\frac{1}{2}}\mu^{-\frac{1}{2}}].$$

$$[K^{-\frac{1}{2}}\mu^{-\frac{1}{2}}]=[LT^{-1}].$$

This shows that $\frac{1}{\sqrt{K\mu}}$ is of the dimensions of a velocity. If we are using the dimensional equations simply to deduce the *dimensions* of any quantity expressed in the one system from its dimensions expressed in the other system, then the above relation is sufficient. If, however, we require to find the *numerical equivalent* for an electrical or magnetic quantity expressed in the one system as expressed in the other, we require to know the numerical value of the ratio $K^{-\frac{1}{2}}\mu^{-\frac{1}{2}}/LT^{-1}$. The value of this ratio can be obtained experimentally by comparing the value of, say, the same quantity of electricity as measured on the two systems. Suppose that a certain quantity of electricity is equal to n_s electro-static units, that is, can be represented in this system by n_s[cm.$^{\frac{3}{2}}$ gram.$^{\frac{1}{2}}$ sec.$^{-1}$ $K^{\frac{1}{2}}$], where K is supposed to be measured in the *c.g.s.* system. Next suppose that this same quantity of electricity, when measured in the electro-magnetic system, is equal to n_m units, or can be represented by n_m[cm.$^{\frac{1}{2}}$ gram.$^{\frac{1}{2}}$ $\mu^{-\frac{1}{2}}$], where as before μ is measured in *c.g.s.* units. Equating these two expressions for the same quantity of electricity, which we may do since they are both expressed in *c.g.s.* units, we get—

$$n_s[\text{cm.}^{\frac{3}{2}}\ \text{gram.}^{\frac{1}{2}}\ \text{sec.}^{-1}\ K^{\frac{1}{2}}]=n_m[\text{cm.}^{\frac{1}{2}}\ \text{gram.}^{\frac{1}{2}}\ \mu^{-\frac{1}{2}}].$$

Hence
$$\frac{n_s}{n_m}[LT^{-1}]=[\mu^{-\frac{1}{2}}\ K^{-\frac{1}{2}}].$$

The quantity $\frac{1}{\sqrt{\mu K}}$ thus represents a velocity of n_s/n_m centimetres per second. Experiment has shown that this velocity is equal to the velocity of light, so that, indicating this velocity by v, we have—

$$[\mu^{-\frac{1}{2}}\ K^{-\frac{1}{2}}]/[LT^{-1}]=v,$$

an expression which allows of our converting electrical quantities ex-

pressed in one system into the other, it being remembered that in each case the units of length, mass, and time employed in the two systems must be the same. When these units are the centimetre, the second and the gram v is equal to 3×10^{10} cm./sec. As an example of the application of this method of converting from one system of units into the other, we may take the following problem. A conducting sphere is placed on an insulating stand at a great distance from all other conductors, and has a radius of 10 cm., what is its capacity expressed in electro-magnetic units? In § 464 it was shown that the capacity of such a sphere in electro-static units was numerically equal to the radius. Hence the capacity of the sphere in electro-static units is $10[\text{cm.}\ K]$. If n_m is the value of the capacity in electro-magnetic units, then, since the dimensions of capacity in this system are $[L^{-1}T^2\mu^{-1}]$, we have the following relation :—

$$10[LK]=n_m[L^{-1}T^2\mu^{-1}],$$

or,

$$\begin{aligned} n_m &= 10[L^2T^{-2}\mu K] \\ &= 10[\mu^{-\frac{1}{2}}K^{-\frac{1}{2}}]^{-2}/[LT^{-1}]^{-2} \\ &= 10 \times \frac{1}{v^2} = \frac{10}{9 \times 10^{20}} = 1.11 \times 10^{-20}. \end{aligned}$$

Hence the capacity of the sphere is 1.11×10^{-20} electro-magnetic units of capacity. Since a microfarad is equal to 10^{-15} electro-magnetic units of capacity, the capacity of the sphere is equal to 1.11×10^{-5} microfarads.

538. The Practical System of Electro-magnetic Units.—It will be convenient for the sake of reference to gather together the relations between the *c.g.s.* electro-magnetic units and those on the practical system, and the following table exhibits these relations :—

Quantity.	Name of Practical Unit.	Equivalent in *c.g.s.* Units.
Current	Ampere . .	10^{-1} *c.g.s.* units.
Quantity	Coulomb .	10^{-1} " "
Electromotive force . .	Volt . .	10^{8} " "
Resistance	Ohm . .	10^{9} " "
Capacity	Farad . .	10^{-9} " "
"	Microfarad .	10^{-15} " "
Energy or work . . .	Joule . .	10^{7} ergs.
Power	Watt . .	10^{7} ergs per second.

These practical units, with the exception of the microfarad, are those which would be obtained if 10^9 cm. were taken as the unit of length and 10^{-11} gram as the unit of mass, the unit of time remaining the second. Thus the dimensions of resistance being $[LT^{-1}\mu]$, if we increase the unit

of length 10^9 times, keeping the unit of time the same, we shall increase the unit of resistance 10^9 times; that is, the ohm is 10^9 *c.g.s.* units. The dimensions of current being $[L^{\frac{1}{2}}M^{\frac{1}{2}}T^{-1}\mu^{-\frac{1}{2}}]$, the increase of the unit of length to 10^9 cm. will increase the unit of current $10^{\frac{9}{2}}$ times, while the change in the unit of mass to 10^{-11} grams will reduce the unit of current $10^{\frac{11}{2}}$ times, so that the net result is that the unit of current on the practical system is $10^{(\frac{9}{2}-\frac{11}{2})}$ times, or $1/10$ of the *c.g.s.* unit.

When considering the thermal effects of currents, it is often convenient to express the results in terms of calories. Since one joule is 10^7 ergs and one calorie is equal to 4.189×10^7 ergs, we get that a joule is equal to $10^7 / 4.189 \times 10^7$ calories or 0.2387 calories. The calorie employed in this reduction is the quantity of heat required to raise the temperature of one gram of water through one degree Centigrade at a temperature of 15° C. Since a watt is equal to one joule per second, it is equal to 0.2387 calories per second.

In order to obviate the use of very large or very small numbers, units are sometimes used which are a million times (10^6) as great or one millionth (10^{-6}) of the practical units. These units are indicated by the prefixes mega- and micro- respectively. Thus a megohm is equal to a million ohms or to 10^{15} *c.g.s.* units of resistance. A microfarad is equal to one-millionth of a farad; that is, 10^{-6} farad or 10^{-15} *c.g.s.* units of capacity.

The term milliampere is sometimes used to indicate a current of a thousandth (10^{-3}) of an ampere.

In electrical engineering it is usual to measure activity or power in kilowatts, a kilowatt being 1000 watts, or 10^{10} ergs per second. Since a horse-power is equal to 7.46×10^9 ergs per second, it follows that a kilowatt is equal to 1.341 horse-power.

PART VIII

ELECTROLYSIS, ELECTROMOTIVE FORCE OF CELLS, AND PASSAGE OF ELECTRICITY THROUGH GASES

CHAPTER XV

ELECTROLYSIS

589. Faraday's Law.—We have in the preceding pages seen that when an electric current passes through a metallic conductor a certain quantity of heat will be developed in the conductor; but after the passage of the current, except for changes caused by the rise of temperature produced by the heat developed in this way, there will be no change either in its chemical composition or physical state. In addition to metals, some liquids conduct electricity, and are called electrolytes, and we now proceed to consider what phenomena accompany the passage of a current through these bodies. The magnetic properties of a circuit which consists wholly or in part of electrolytes differ in no way from those of a circuit composed of metals only, and hence do not require any further consideration. The passage of a current through an electrolyte is accompanied, however, not only by the production of heat as in a metallic conductor, but also by certain chemical changes which take place in the electrolyte, and we now proceed to consider these in detail.

When a current is passed through an electrolyte, such as a solution of sulphuric acid in water, by dipping two platinum plates into the solution, and connecting one of these, called the *anode*, with the positive pole of a battery, and the other, called the *kathode*, with the negative pole, decomposition of the electrolyte will accompany the passage of the current. The two products of the decomposition of the electrolyte, whether they are either or both elements or compounds, will be liberated one at the cathode and the other at the anode, and not at all at any point of the liquid between. That part of the electrolyte which is liberated at the anode is called the *anion*, and that part liberated at the kathode the *kation*. It does not necessarily follow that the anion and kation are actually given off as such at the anode and kathode respectively, for secondary chemical changes often take place between the ions and the electrodes, as the plates used to form the anode and kathode

are called, or with the undecomposed portion of the electrolyte. Thus in the case of the electrolysis of a solution of sulphuric acid (H_2SO_4), the kation is H, while the anion is SO_4. But while hydrogen is given off at the kathode, at the anode a secondary reaction takes place, the SO_4 reacting with the water of the solution so as to produce sulphuric acid and free oxygen according to the equation

$$SO_4 + H_2O = H_2SO_4 + O.$$

The laws which govern electrolysis were discovered by Faraday, and are hence known as Faraday's laws. These are :—

1. The quantity of an electrolyte decomposed is proportional to the quantity of electricity which passes.

2. The mass of any ion liberated by a given quantity of electricity is proportional to the chemical equivalent weight of the ion.

In the case of elementary ions the chemical equivalent weight is the atomic weight divided by the valency, while in that of a compound ion it is the molecular weight divided by the valency. If the weight of an ion liberated by the passage of the unit quantity of electricity is called the electro-chemical equivalent of the ion, then Faraday's laws can be put into the form :—

The mass of an ion liberated is equal to the product of the quantity of electricity which passes into the electro-chemical equivalent of the ion ; the electro-chemical equivalents of the ions being to one another as the chemical combining weights of these ions.

Since, if the electro-chemical equivalent of any one ion is known, that of any other can be calculated from the chemical equivalent weights, it is of importance to determine the value of the electro-chemical equivalent in the case of one ion. Accurate experiments have shown that when one coulomb of electricity passes, that is, when a current of one ampere passes for one second, the weight of silver deposited from a solution of a silver salt is 0.001118 grams. Since the atomic weight of silver is 107.94, and the valency is 1, while the atomic weight of hydrogen is 1, and its valency is also 1, the electro-chemical equivalent of hydrogen is equal to 0.001118/107.94, or .000010357. Hence a current of A amperes flowing for t seconds will liberate $1.0357 \times 10^{-5} At$ grams of hydrogen, or, if q is the chemical equivalent weight of any ion, will liberate m grams of this ion where m is given by

$$m = 1.0357 \times 10^{-5} qAt.$$

As an example, in the case of copper as a cupric salt, the atomic weight is 63, while the valency is 2 ; hence the chemical equivalent is 63/2, and the electro-chemical equivalent of copper is $1.0357 \times 10^{-5} \times 31.5$.

Since the passage of 1 coulomb will deposit .001118 grams of silver, it will require the passage of 107.94/.001118 coulombs, or 96,550 coulombs, to deposit 1 gram equivalent, that is, the chemical equivalent in grams, of

silver. By Faraday's second law it follows that the passage of 96,550 coulombs will cause the separation of one gram equivalent of any kind of ion. In the case of ions which can have more than one chemical valency there will be more than one chemical equivalent. Thus iron can exist in a compound either in the ferric condition, when it has a valency 3, and consequently a chemical equivalent of 56/3 or 18.7, or as a ferrous salt, when it has a valency of 2, and hence the chemical equivalent is 56/2, or 28. Thus when a ferric salt is electrolysed the electro-chemical equivalent of iron is $1.0357 \times 10^{-6} \times 18.7$; while when a ferrous salt is electrolysed the electro-chemical equivalent is $1.0357 \times 10^{-6} \times 28$.

Since the passage of 96,550 coulombs of electricity through an electrolyte always liberates one gram equivalent of each ion, if we suppose the electricity to pass by a kind of convection, being carried by the ions, a positive charge being carried by the kations in the direction of the current, and a negative charge by the anions in the opposite direction, it follows that the charge carried by the chemical equivalent of each ion must be the same. In the case of univalent ions the electro-chemical equivalents are proportional to the atomic or molecular weights according as the ion is an element or a compound. Hence, if we extend the term ion to mean the smallest portion of the substance producing the ion which can take part in a chemical reaction, the charge carried by each ion must, in the case of all univalent ions, be the same. Let ϵ be the charge carried by a univalent ion, then if w is the weight of the ion of hydrogen, one gram of hydrogen will correspond to $1/w$ ions, and since the quantity of electricity transported by one gram of hydrogen ions is 96,550 coulombs, the quantity transported by each ion is given by

$$\epsilon = 96550 \times w,$$

or, if we suppose that an ion of hydrogen is the same as the atom, and that an atom of hydrogen weighs 8.3×10^{-25} grams,

$$\epsilon = 96550 \times 8.3 \times 10^{-25} = 8 \times 10^{-20} \text{ coulomb.}$$

Thus the quantity of electricity transported by a univalent ion is 8×10^{-20} coulomb. If the ion is a kation ϵ is positive, that is, each ion carries a positive charge. If, however, the ion is an anion, the charge is negative, and is transported in the opposite direction to that in which the current flows.

In the case of a divalent ion, such as copper or SO_4, the charge carried by each ion must be equal to $\pm 2\epsilon$, for each atom of copper weighs 63 times as much as an atom of hydrogen, while the weight of copper deposited by the passage of a given quantity of electricity is only 63/2 times as much as the weight of hydrogen liberated by the same quantity of electricity. Thus the number of ions of copper deposited by one coulomb is half the number of hydrogen ions liberated by the same quantity of electricity, and hence each copper ion must carry twice as

great a charge as each hydrogen ion. In the same way the charge carried by trivalent ions, such as aluminium, must be $\pm 3\epsilon$.

540. Electrolytic Dissociation.—Careful experiments have shown that in the case of electrolytes Ohm's law holds, that is, an electromotive force acting in the electrolyte produces a current which is proportional to the E.M.F. As we shall see later (§ 544), when considering the E.M.F. which is acting to produce a current through an electrolyte, it does not do to measure the E.M.F. between the electrodes by means of which the current is conveyed to and from the electrolyte, for it requires in general a definite E.M.F. to cause a current to pass from a metal to an electrolyte, so that, when considering the connection between the E.M.F. and the current which it produces, that is, the question of the resistance of electrolytes, the difference of potential must be measured between two points *within* the electrolyte itself.

Ohm's law being true for electrolytes, it follows that Joule's law must also be true, and hence all the energy of the current spent when traversing an electrolyte must be used simply in the production of heat, and none of it can be employed in doing chemical work in splitting up the electrolyte into ions. Within the mass of the electrolyte, therefore, the action of the current in electrolysis must simply consist in the exertion of a directive influence on the charged ions, causing them to move towards the electrodes, where, as we shall see, the work corresponding to the splitting up of the chemical compound is performed. When seeking the explanation of how, at any rate, a part of the electrolyte can be in such a condition as to allow the anions and kations to be moved in opposite directions, we are at once met with the curious fact that it has been proved experimentally that perfectly pure water is practically a non-conductor, as is also gaseous hydrochloric acid, while a solution of hydrochloric acid in water conducts freely. In the same way pure sulphuric acid is a non-conductor, or at any rate a very bad conductor, while a dilute solution of sulphuric acid is a comparatively good conductor. If we suppose that the ionisation, as the process which consists in so changing the relations of the constituents of a compound that they are able to conduct electricity is called, is due to the shaking apart of the ions in the compound molecule during the collisions between two molecules, then we should expect that the more frequent the collisions the greater the proportion of the molecules which are ionised, and hence the greater the electrical conductivity. As we have mentioned, however, this is not the case, for pure sulphuric acid does not conduct.

It is therefore evident that the hydrochloric acid and the sulphuric acid when they are dissolved in water are in a different condition from that in which they were before solution. In fact, in the solution a greater or less proportion of the molecules are either permanently split up into their ions, so that in the place of a molecule, say, of HCl we have a hydrogen ion with its positive charge $+\epsilon$, and a chlorine ion with

its negative charge $-\epsilon$, or the forces which bind the H and Cl ions in the molecule of HCl are so reduced that during the collisions which occur these ions become separated, so that at any instant a finite proportion of ions are separated one from the other. It does not follow that the same ions are always thus separated, as they may recombine, but that taken as a whole there are always a considerable number in the separate conditions. In either case the HCl is said to be dissociated or ionised by its solution in the water. In very weak solutions the electrical resistance is such that it would appear that all the molecules of the HCl are dissociated, while in stronger solutions only a fraction of the HCl molecules are dissociated, and that the undissociated molecules play no part in the conduction of the electricity in the solution.

In order to account for the dissociating influence of water, the theory has been put forward that the forces which hold the ions together to form a molecule are due to the electrical attractions between the oppositely charged ions, so that as the specific inductive capacity of water is very great, this force is very much reduced when the molecule is dissolved in water. For, as we have already seen (§ 462), if two charged bodies are transferred from air into a medium of which the specific inductive capacity is K, the force exerted between them is reduced in the proportion of K to 1. Hence, since the ions are supposed to have a constant charge, the force exerted between the ions in a molecule, tending to prevent the splitting up of the molecule, will be less in a medium of high specific inductive capacity such as water, for which $K=79$, than in one of small specific inductive capacity.

The dissociation or ionisation here considered is of a different nature from that which may be produced by increasing the temperature to which a pure substance is subjected, and must not be confused with it. Thus ammonium chloride (NH_4Cl) when heated dissociates into ammonia (NH_3) and hydrochloric acid (HCl), while in a solution of NH_4Cl in water dissociation takes place into the ions NH_4 and Cl.

The hypothesis that in an electrolyte the ions exist in the uncombined condition receives further support from the other properties of such solutions considered in §§ 165, 225, 227. As has been pointed out, in the case of solutions of acids and alkalies, the osmotic pressure, the lowering of the freezing-point, and the lowering of the vapour pressure are in all cases much greater than in the case of non-electrolytes. If we remember that on Van't Hoff's hypothesis these effects are proportional to the number of molecules present, then, in order to extend this hypothesis to electrolytes, we are led to postulate the presence of a greater number of molecules than appears to be present, if we suppose that the body is not dissociated. Thus, as shown on page 268, the molecular depression of the freezing-point produced in dilute aqueous solutions of such non-electrolytes as sugar, glycerine, acetic acid, and ethyl alcohol, is about 19. In the case of hydrochloric acid, sulphuric acid, and sodium chloride,

the molecular depression is about 36. Now if in the dilute solutions the molecule of hydrochloric acid is dissociated into H and Cl ions, and if each ion produces the same effect as an undissociated molecule, the solution will contain twice as many active molecules as would be the case if no such dissociation occurred, and hence we should expect the depression of the freezing-point to be twice as great as that in the case where no dissociation takes place. In the case of a substance like barium chloride ($BaCl_2$), which can dissociate into a barium ion and two chlorine ions, we should, on the dissociation hypothesis, expect the molecular depression of the freezing-point to be three times as great as in the case of an undissociated body. As a matter of fact, the value obtained in the case of barium chloride is 48.6, and although this is not quite equal to three times 19, yet the difference can be satisfactorily explained by supposing that the whole of the barium chloride in the solution employed in determining the depression of the freezing-point was not dissociated, so that there were some molecules of $BaCl_2$ present which would only produce a third of the depression that they would have produced had they been dissociated.

While some supporters of the dissociation theory maintain that in a dilute solution the ions are almost, if not quite, separate and remain so all the time, others only consider that in an electrolyte the ions are so loosely joined together that they are continually exchanging partners, and that the influence of the electromotive force used to send a current is to direct those ions which at the moment happen to be in the process of changing partners. Several attempts have been made to devise experiments to show that the ions are separate, but they are none of them conclusive. It does not, however, much matter which hypothesis we adopt, since either is capable of explaining the observed phenomena.

541. Migration of the Ions.—If a solution of copper sulphate is electrolysed between copper electrodes, the copper kation will be separated from the solution and be deposited on the kathode. The anion SO_4 will not, however, be liberated in the free state at the anode, but will attack the anode forming copper sulphate. The result is that for every copper ion that goes out of solution at the kathode, another copper ion comes into solution at the anode, while the SO_4 ions remain in a constant number in the solution. Thus the total quantity of copper sulphate in the solution is unaltered by the process of electrolysis. The concentration of the solution, however, becomes greater near the anode, while it becomes less near the kathode. In the case of a solution of copper sulphate this can be easily seen, if the electrodes are arranged one above the other, the anode being below, so that the strength of the solution does not tend to be equalised by convection currents set up by the differences in density of the solution near the electrodes. After the current has been passed for some time, it will be observed that the colour

of the solution is darker near the anode and lighter near the kathode than it was before the passage of the current.

In this case the reaction between the anion, SO_4, and the anode to produce $CuSO_4$ will account for the strengthening of the solution near the anode. This explanation will not, however, hold in all cases, for Hittorf, who made many accurate observations of the changes in the strength of electrolytes produced by the passage of the current, found, in the case of many electrolytes where no secondary reaction between the liberated ions and either the electrodes or the solvent took place, that the solution became weaker near one electrode than near the other. In such cases as that of the electrolysis of copper sulphate, where secondary reactions take place, it is found that, when due allowance is made for the effects of such secondary reactions on the concentration of the solution near the electrode at which the secondary reaction occurs, still, the concentration varies more near one electrode than near the other. In order to account for these facts, Hittorf has put forward the hypothesis that the ions, when they travel through the liquid under the influence of the potential difference which forces the current through the solution, do not travel with the same velocity, but that for a given concentration of the solution, and a given potential gradient, each kind of ion moves with a constant velocity, and that the velocities corresponding to different ions are not the same. Hittorf also found that the weakening of the solution took place exclusively in the immediate neighbourhood of the electrodes.

Let a be the weight of salt lost by the solution near the anode during the passage of a given quantity of electricity, and k the corresponding quantity near the kathode. Then Hittorf found that the ratio a/k has a fixed value for every electrolyte, if only the solution is very dilute. The total quantity of salt lost from the solution is $a+k$, and the ratio n of the loss of salt near the kathode to the total loss, or $k/(a+k)$, is called the migration constant or transport number of the *anion*. In the same way, $a/(a+k)$ or $1-n$ is called the migration constant or transport number of the *kation*.

Assuming that the differences in the concentration are due to the fact that the ions move or migrate with different velocities, we can deduce the relative velocities of the ions in the case of any electrolyte if we know the value of n. Thus suppose that a dilute solution of hydrochloric acid is electrolysed, the process being continued till 96,550 coulombs of electricity have passed, so that 36.4 grams of the acid are decomposed, then one gram of hydrogen will be liberated at the kathode, and 35.4 grams of chlorine at the anode.

In the first place, let us suppose that the current is carried from the anode, A (Fig. 517), to the kathode, K, exclusively by the motion of the hydrogen ions, each carrying its charge $+\epsilon$. If then the hydrogen ions are represented by the sign $+$, and the chlorine ions by $-$, the arrangement of the ions, or, if we like, the molecules of salt, in their loosely

combined condition can be represented diagrammatically by (*a*), Fig. 517. During the passage of the current, all the H ions will move to the left, so that 1 gram of hydrogen ions will be evolved at K, and the solution near A will therefore lose 1 gram of H ions. The solution will also lose 35.4 grams of Cl ions near A, for the Cl ions which are left on account of the migration of the H ions will be liberated at the anode. Thus the solution near A will lose 36.4 grams of HCl, so that the arrangement of the ions or molecules after the passage of the current can be represented by (*b*). Hence the supposition that the H ions are the only ones which migrate, necessitates the whole loss of concentration of the electrolyte occurring at the anode. In the same way, if we suppose that the Cl ions are the only ones which convey the current, each moving towards the anode with the charge $-\epsilon$, the H ions remaining distributed uniformly throughout the solution, and not moving when the current passes, then the solution near the kathode will lose the whole 36.4 grams of HCl.

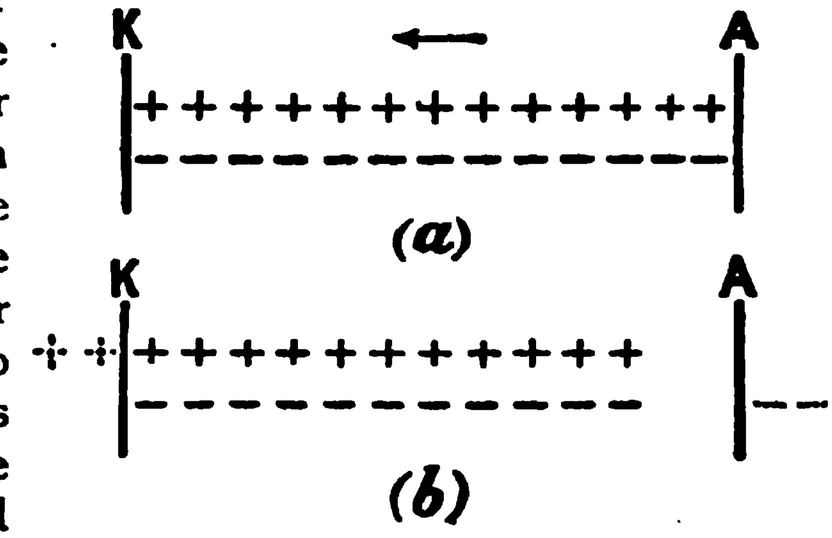


FIG. 517.

Next suppose that the H ions move with a velocity u, and the Cl ions with a velocity v, then part of the current will be due to the movement of the positively charged H ions in the same direction as the current, and the remainder will be due to the movement of the negatively charged Cl ions in the opposite direction. In unit time the number of H ions which reach the kathode will be to the number of Cl ions which reach the anode as $u:v$. Hence the fraction of the quantity of electricity which passes, which is due to the movement of the H ions, is $u/(u+v)$, while the fraction due to the movement of the Cl ions is $v/(u+v)$. In order to obtain the changes in concentration of the solution at the two electrodes, we have therefore to calculate what quantity of the electrolyte will be lost near the kathode due to the passage of $\frac{v}{u+v}\times 96550$ coulombs carried by the Cl ions, and the loss near the anode due to the carrying of $\frac{u}{u+v}\times 96550$ coulombs by the H ions. From the considerations of what happened when the whole of the electricity was supposed to be carried by the H ions, we see that if $\frac{u}{u+v}\times 96550$ coulombs is carried by the H ions, the loss (a) of HCl near the anode will be $36.4\times\frac{u}{u+v}$ grams. In the same way, the loss (k) near the kathode will be $36.4\times\frac{v}{u+v}$ grams.

Hence $\frac{u}{v}=\frac{a}{k}=\frac{1-n}{n}$, where n is the migration constant for the anion, as determined by Hittorf. It will be noted that in unit time u hydrogen ions move away from the anode, and hence u chlorine ions will be left alone near the anode, and will be evolved. In the same way, when v Cl ions move away from the kathode, v H ions will be left, and these also will be liberated, so that at both anode and kathode $u+v$ ions will be liberated.

542. The Molecular Conductivity of Electrolytes—Ionic Velocities.—Let c be the number of gram equivalents (§ 539) of an electrolyte contained in a litre of solution, then the actual number of molecules in a cube of which each edge is one centimetre will be proportional to c. Hence, if the whole number of molecules of the salt are dissociated, the number of ions contained in the unit cube of the solution will be proportional to c. If now an E.M.F. e acts between the opposite faces of the unit cube, and a current C passes, and if K is the conductivity (§ 481) of the solution, we have $K=C/e$. If further we assume that the current is conveyed exclusively by the ions, and that, as long as the rate of fall of potential along the solution is constant, that is, in the case before us if e is constant, the velocity of the ions is constant whatever the concentration of the solution, it follows that the current must be proportional to the number of ions between the opposite faces of our unit cube. Hence the conductivity must be proportional to the number of ions contained in the unit cube, so that if K is the specific conductivity of the solution, that is, the conductivity between the opposite faces of the unit cube, we have

$$K=mc,$$

where m is a constant, and is called the molecular conductivity of the salt which forms the electrolyte.

Kohlrausch, who has made a number of measurements of the conductivity of solutions of electrolytes of different concentrations, found that for fairly strong solutions the molecular conductivity was not constant, but that it increased as the dilution increased. As the solution became very dilute, however, the value obtained for the molecular conductivity for any given salt, in general, became constant. Thus in the case of solutions of potassium chloride the following table exhibits the values which he obtained.

The first column contains the value of the concentration, that is, the number of gram equivalents of the salt contained in a litre of the solution (1 gram equivalent of KCl=74.4 grams). The second column contains the specific conductivity of the solution at 18° C., measured in ohms^{-1} cm.^{-1}, while the third column contains the values of the molecular conductivity deduced from the values given in the first two columns.

CONDUCTIVITY OF SOLUTIONS OF KCl.

Concentration in Gram Equivalents of KCl per Litre.	Specific Conductivity in $Ohms^{-1}$ $cm.^{-1}$	Molecular Conductivity.
3.0	.2637	.0879
1.0	.0977	.0977
.5	.0509	.1018
.1	.01113	.1113
.01	.001219	.1219
.001	.0001268	.1268
.0001	.00001295	.1285
.00001	.000001293	.1293

This table very clearly shows the increase of the molecular conductivity as the dilution is increased, and it will be noticed that, while the conductivity of a solution containing one gram equivalent in the litre is 0.0977, if a solution of one-tenth of a gram equivalent per litre is taken, the conductivity is 0.01113. Hence, while in the second case there is only a tenth of the number of molecules in the solution to conduct the current, the conductivity, instead of being a tenth of that at the greater concentration, is 0.114. It would thus appear that as the concentration decreases the salt conducts better. This change is explained on the ionic hypothesis by supposing that, at the greater concentration, only part of the total number of molecules of the salt present in the solution is dissociated into ions, and so the number of ions capable of conveying the current is less than the number calculated on the supposition that all the salt is dissociated, and further that the proportion of salt dissociated increases as the dilution is increased. As the dilution is increased, the molecular conductivity increases, and for very great dilution becomes practically constant, owing to the fact that at great dilutions the whole of the salt present is dissociated. Under these circumstances, when calculating the molecular conductivity we do not, as is the case at greater concentrations, include the undissociated molecules, which, although they are reckoned in the value taken for the concentration, c, are not effective in the current conduction.

By a study of the molecular conductivity of very dilute solutions of different salts containing a common anion, such as potassium chloride and sodium chloride, and of similar solutions of salts containing the same kation, such as sodium chloride and sodium nitrate, Kohlrausch was led to the conclusion that the maximum molecular conductivity of an electrolyte can be calculated by adding together two constants, the values of which depend on the nature of the anion and kation respectively. Further, that these two constants for the ions of any given

electrolyte must be proportional to the migration velocities u and v of the ions.

To see how this consequence follows from the ionic theory, we may consider a column of the electrolyte of which the cross-section is one square centimetre, and suppose that the difference of potential between the ends of a length of this cylinder of one centimetre is e volts, the specific conductivity being K. The current C flowing through the liquid will by Ohm's law be given by $C=eK$. Now if we consider a partition across the column of electrolyte, and if u is the migration velocity of the kation and v that of the anion, and if N is the number of anions and of kations, respectively, contained in the unit of volume, the number of anions which cross this partition in unit time is Nv, while the number of kations which cross in the opposite direction is Nu. Since each kation carries a charge of $+\epsilon$, and each anion one of $-\epsilon$, the total quantity of electricity carried across the partition in one second is $(u+v)N\epsilon$, for a charge $-\epsilon Nv$ carried in the opposite direction to that in which the current is flowing is the same as the passage of $+\epsilon Nv$ in the opposite direction. But the total quantity of electricity which crosses the partition in one second is the same thing as the current C which is passing through the liquid. Hence

$$C=(u+v)N\epsilon=eK.$$

Now $N\epsilon$ is the total charge on all the ions of one sign within one cubic centimetre of the solution, and we have seen on p. 787 that the total charge on all the ions corresponding to one gram equivalent is 96,550 coulombs or 9655 *c.g.s.* units. Hence, if the solution we are considering contains c gram equivalents per litre, that is, $c \times 10^{-3}$ gram equivalents per cubic centimetre, the total charge on the ions of one sign in 1 c.c. will be $9655 \times c \times 10^{-3}$, so that $N\epsilon = 9655 \times c \times 10^{-3}$. Hence

$$u+v=\frac{eK}{9655c} \times 10^3.$$

If the potential gradient is one volt per centimetre, so that e is one volt or 10^8 *c.g.s.* units, we get—

$$u+v=1.0357 \times 10^7 \frac{K}{c}.$$

But K/c is the molecular conductivity m. Hence

$$u+v=1.0357 \times 10^7 \,.\, m,$$

where m is measured in *c.g.s.* units. If m' is the molecular conductivity measured in $\text{ohms}^{-1}\ \text{cm.}^{-1}$, since 1 ohm $=10^9$ *c.g.s.* units—

$$u+v=1.0357 \times 10^{-2} m'.$$

Now we have already seen that the migration constant of the anion was connected with the migration velocities by the equation

$$n = \frac{v}{u+v}.$$

Hence, knowing from migration data the value of the ratio of the velocity of the anion to the sum of the velocities of the two ions, and from the conductivity data the value of the sum of the two velocities, the absolute value of each of these quantities can be calculated. In the following table a few of the values of the migration velocities obtained in this way are given, as well as the data from which they are calculated. The temperature is 18° C. and the potential gradient 1 volt per cm. The first column contains the name of the electrolyte, the second the molecular conductivity for an infinitely dilute solution, in which all the salt may be considered to be dissociated, the third column contains the sum of the ionic velocities obtained by multiplying the numbers in the second column by 1.0357×10^{-2}. The fourth column contains the numbers obtained by Hittorf for the ratio of the loss of salt at the kathode to the total loss, that is, the values of n, while in the last two columns the values of the velocities of the kations and anions are given as derived from the numbers in the third and fourth columns.

MIGRATION VELOCITY OF THE IONS.

Substance.	m_∞.	$u+v$ cm. per sec.	$\frac{v}{u+v}$	u cm. per sec.	v cm. per sec.
KCl . . .	1307×10^{-4}	135×10^{-5}	.51	66×10^{-5}	69×10^{-5}
NaCl . .	1095 "	113 "	.62	43 "	70 "
LiCl . . .	1010 "	105 "	.68	34 "	71 "
NH_4Cl . .	1297 "	134 "	.51	71 "	63 "
HCl . . .	3752 "	389 "	.21	311 "	78 "
KNO_3 . .	1254 "	130 "	.50	65 "	65 "
$NaNO_3$.	1042 "	108 "	.61	42 "	66 "
$AgNO_3$. .	1159 "	120 "	.53	56 "	64 "
KOH . .	2360 "	244 "	.74	63 "	181 "
NaOH . .	2137 "	221 "	.84	35 "	186 "

It will be seen that the numbers obtained for the same ion from different salts agree fairly well together, the differences being probably due to inaccuracies in the values for the migration data. This is rendered extremely probable on account of the difficulty of accurately measuring the loss of salt near the electrodes. Further, it is to be noted that the migration data have only been determined for comparatively concentrated solutions, and that the values of the migration constants seem to change slightly with the concentration, so that, instead of using the values given in the table, we ought to use those for a very dilute solution, such as is

employed in driving the molecular conductivity. Kohlrausch has attempted to allow for this effect of concentration on the migration constant, and the values for the ionic velocities which he considers to be the most accurate are given in the following table :—

VELOCITY OF THE IONS.

Kations.		Anions.	
K . . .	66×10^{-6} cm. per sec.	Cl . . .	69×10^{-6} cm. per sec.
Na . . .	45 " "	NO_3 . .	64 " "
Li . . .	36 " "	OH . . .	182 " "
H . . .	320 " "		
Ag . . .	57 " "		

Direct experimental measurements have been made of the velocities of some ions, and the results obtained agree with those given in the above table, as calculated from migration and conductivity data.

543. The Ionisation Coefficient.—On the dissociation theory, the explanation of the increase of the molecular conductivity with the dilution is that more of the molecules of the salt become dissociated into ions as the dilution increases. If a is the fraction of the total number of molecules of the salt present which have become dissociated at the given concentration, then, since it is only the dissociated molecules which can conduct, it follows that the molecular conductivity at any given concentration must be to the molecular conductivity at infinite dilution, when all the molecules are dissociated, as the number of molecules actually dissociated in the solution considered is to the total number of molecules in the solution. Hence, if m_c is the molecular conductivity when the concentration is c and m_∞, the molecular conductivity for an infinitely dilute solution,

$$a = m_c \,/\, m_\infty.$$

The quantity a, which expresses on the dissociation theory the fraction of the number of molecules present which exist in the solution in the ionic condition, is called the ionisation coefficient or the dissociation coefficient.

Now according to Van't Hoff's theory of solutions the osmotic pressure (§ 165) of a given solution is proportional to the number of molecules present in the solution. Suppose that N molecules of potassium chloride are dissolved in 1 c.c. of water, then, if none of the molecules become dissociated, the osmotic pressure ought to be proportional to N. If, however, n of the molecules of KCl become dissociated into the ions K and Cl, the number of molecules present will be $(N-n)$ undissociated molecules of KCl together with $2n$ ions. Hence, if we suppose that the effect of each of the ions, into which the molecule is split, in producing the osmotic pressure is the same as that produced by one of the undissociated molecules, the osmotic pressure will be proportional to $N+n$. Thus,

by determining first the osmotic pressure, p, produced when N molecules of some body, such as sugar, which is not an electrolyte and hence is not dissociated, is dissolved in unit volume of water, and, secondly, the osmotic pressure, p', produced when the same number of molecules of an electrolyte is dissolved, we can calculate the ratio of the number of molecules, n, which are dissociated to the total number of molecules present, that is, calculate the ionisation coefficient a. Thus

$$\frac{p'}{p}=\frac{N+n}{N}=1+\frac{n}{N}=1+a,$$

or

$$a=\frac{p'-p}{p}.$$

In the same way, from the comparison of the depression of the freezing-point (§ 225) produced by N molecules of a non-electrolyte with that produced by the same number of molecules of an electrolyte we can, making the same supposition as to each of the ions of a dissociated molecule producing the same depression as an undissociated molecule, calculate from such observations the value of the ionisation coefficient.

In the following table the values of the ionisation coefficients as obtained from these entirely different data are given, and it will be noticed that the agreement of the numbers obtained in the various ways is on the whole very fair. Whether we accept the dissociation hypothesis or not, at any rate these numbers show that there must be some intimate relation between the cause of the conductivity of electrolytes, the depression of the freezing-point and the osmotic pressure.

IONISATION COEFFICIENTS.

Substance.	Concentration in Gram-Molecules per Litre.	Coefficient of Ionisation Deduced from Conductivity.	from Depression of Freezing-Point.
HCl	0.002	1.00	0.98
	0.01	0.99	0.96
	0.1	0.94	0.89
H_2SO_4	0.003	0.90	0.86
	0.005	0.85	0.84
	0.05	0.62	0.61
KOH	0.002	1.00	0.98
	0.01	0.99	0.94
	0.1	0.93	0.83
$NaOH$	0.002	0.99	0.98
	0.01	0.99	0.94
	0.05	0.90	0.88
Na_2CO_3	0.003	0.91	0.96
	0.005	0.86	0.96
	0.05	0.65	0.73

544. Polarisation.—If an electrolytic cell is prepared containing a solution of sulphuric acid, the electrodes being composed of platinum plates, and this cell and a galvanometer are included in a circuit together with a source of E.M.F., the following phenomena will occur. If the E.M.F. of the battery is less than about 1.7 volts, on closing the circuit the galvanometer will indicate that a current passes in the circuit. The strength of the current will, however, rapidly decline, till after a short time only a very minute current will continue to pass. If the E.M.F. of the battery is greater than 1.7 volts the current will decrease in strength for some time after the closing of the circuit, but it will never become evanescent, as was the case when the E.M.F. was below 1.7 volts. If, after the passage of a current through the electrolytic cell, the battery is removed and the circuit completed by joining together the ends of the wires which were connected to the poles of the battery, a current will pass round the circuit for some time in the opposite direction to that in which the current sent from the battery passed. Thus the plates of platinum immersed in the electrolyte possess the power not only of practically stopping the passage of a current in a circuit in which the E.M.F. is less than 1.7 volts, that is, are capable of exerting an E.M.F. in the opposite direction to that which is due to the battery, and so of preventing the passage of a current; but also this opposing E.M.F. continues for some time after the removal of the external E.M.F., so that the electrodes are able to send a current through the circuit. This phenomenon is called polarisation, and the electrodes are said to be polarised. If a current is passed through an electrolytic cell containing dilute sulphuric acid, and in which the electrodes are of platinum, and the E.M.F. between the electrodes is measured immediately after the removal of the external E.M.F., it will be found to be 1.07 volts, the anode being at the higher potential.

It must be noticed that, although according to the ionic hypothesis the ions exist in the electrolyte in the dissociated condition, it does not follow that no work has to be done to liberate the ions from the solution. In the solution each ion has its appropriate charge; when the ion is liberated at the electrode, however, this charge has been removed, so that the condition of the liberated ions is quite different from that when they were in the solution. If we assume that the hydrogen, say, as it is given off, consists of molecules each containing two atoms, these atoms being held together by chemical forces so as to form a compound, containing, however, only one kind of element; then if, as seems probable, chemical combination really consists in the holding together of the atoms by the electrical forces in play between their charges, we are led to the necessity for supposing that the molecule of hydrogen given off at the kathode must consist of a positively charged atom and a negatively charged atom. Hence in the process of electrolysis, while one hydrogen atom retains its positive charge the other loses its positive charge, and takes up from

the kathode an equal negative charge, and the two combine to form a molecule of neutral hydrogen. As far as the passage of electricity through the electrolytic cell is concerned, the giving up of its positive charge to the kathode by a hydrogen ion is exactly the same thing as taking an equal negative charge from the kathode, so that although all the hydrogen ions do not lose their charge when they are liberated, as we have tacitly assumed in the preceding pages, the quantity of electricity which passes through the cell, while a given number of H ions are liberated, is the same as it would be if all the H ions gave up all their charges to the kathode.

The explanation of the fact that the polarised electrodes are able, when the E.M.F. sending a current through the cell is removed, to send a current through the circuit in the reverse direction, is that the gases produced at the electrodes are not entirely liberated and given off, but that the platinum absorbs a certain quantity of the gas. On the removal of the E.M.F. this absorbed gas tends to return into the solution, and each ion of the kation, when it leaves the kathode, becomes charged positively, while each anion as it leaves the anode is negatively charged. Thus positive electricity is taken away from the kathode and negative from the anode, and hence in the external circuit connecting the electrodes a current will flow from the anode to the kathode, that is, in the reverse direction to the original current.

There are two distinct effects which are in general included under the term polarisation. One of these is the back E.M.F., which must exist when chemical decomposition is being performed by the current, in order that the requisite amount of energy may be supplied by the current. The other is an effect due to the accumulation of the products of the decomposition on or near the electrodes. Thus in the case of the electrolysis of dilute sulphuric acid between unplatinised platinum electrodes, the polarisation E.M.F. amounts to about 1.7 volts. Le Blanc has, however, shown that if platinised electrodes are employed, water may be decomposed with an E.M.F. of only 1.07 volts. When unplatinised electrodes are used the gases are evolved in the form of bubbles which form on the plates, and it would appear that a certain amount of work has to be done in producing these bubbles. With platinised electrodes, on the other hand, a much larger quantity of the gases separated by the passage of the current will not be liberated in the gaseous form, but will be absorbed by the platinum, and, in addition, the numerous small points which are present on the platinised surface seem to facilitate the evolution of the bubbles. Hence we are led to the conclusion that 1.07 volts represents the E.M.F., which corresponds to the chemical work,[1] that is, the splitting up of the chemical compound that forms the electrolyte, which is done in the cell, while the greater value, 1.7 volts, which is necessary to produce decomposition when unplatinised platinum electrodes are employed, is

[1] We shall return to the subject of the connection between the E.M.F. and the energy required to perform the chemical work in § 558.

due to the additional work which has to be done at the electrodes in performing the mechanical work of separating the liberated gases from the electrodes.

The term polarisation is also used in yet another sense, when the increase of the resistance of an electrolytic cell, due to the fact that the bubbles of gas which adhere to the electrodes virtually diminish the cross-section of the liquid conductor, and hence increase the resistance, is spoken of as being due to polarisation.

It is better, however, to call such an effect as this a transition or secondary resistance. The magnitude of these mechanical effects depends on the current density, that is, the quotient of the current passing by the area of the electrode, on the solubility of the gases liberated in the electrolyte, and on the occlusion of the gases by the electrodes. So that if the term polarisation is taken to include all these effects its value is most indefinite.

In the case of the electrolysis of a solution of copper sulphate between electrodes of copper we have copper deposited on the kathode, and the SO_4 ion attacks the anode producing copper sulphate. Hence in such a case the energy which is required to split up the salt is regained by the formation of an equal mass of the salt at the anode, so that no work has to be done by the current in producing chemical energy of separation. We should therefore expect that in this case there would be no counter E.M.F. of polarisation, which as a matter of fact is found to be the case. Similarly, when a solution of zinc sulphate is electrolysed between electrodes of zinc there is no polarisation. When there is no polarisation, the difference of potential between the electrodes during the passage of a current is equal to the product of the current into the resistance between the electrodes according to Ohm's law. If, however, there is an opposing E.M.F. of polarisation e developed when a current C is passed, the resistance of the electrolyte between the electrodes being R, then by Ohm's law there will be a difference of potential between the portions of the electrolyte near the anode and kathode respectively given by RC. Hence in this case the difference of potential between the *electrodes* will be given by $E=RC+e$, so that $C=(E-e)/R$. Hence, when we are considering the passage of a current between the electrodes in a cell where polarisation occurs, the applied E.M.F. must be reduced by the E.M.F. of polarisation in order to calculate the current, according to Ohm's law, from the resistance of the cell.

The measurement of the polarisation produced by a given current must be made very quickly after the removal of the external E.M.F. which was sending the current, for the magnitude of the polarisation E.M.F. falls rapidly. The experiment may be performed by means of the arrangement shown in Fig. 518. The electrodes of the electrolytic cell C are connected, one with the stem of an insulated tuning-fork F, and the other with one pole of a battery, and with one of the pairs of quadrants of a quadrant electrometer. The other pole of the battery and

the other quadrants of the electrometer are connected to two small mercury-cups, *a* and *b*. Two platinum wires are attached to the prongs of the fork in such a way that when the prongs move towards one another one wire dips into the mercury-cup *b*, and connects the electrolytic cell with the battery, while when the prongs move away from each other the wire attached to one prong is withdrawn from the cup *b*, while that attached to the other is dipped into the cup *a*, and hence connects the electrolytic cell with the electrometer. Thus as the fork vibrates the electrolytic cell is alternately connected with the battery and with the electrometer, and from the deflection of the electrometer the E.M.F. of polarisation can be obtained.

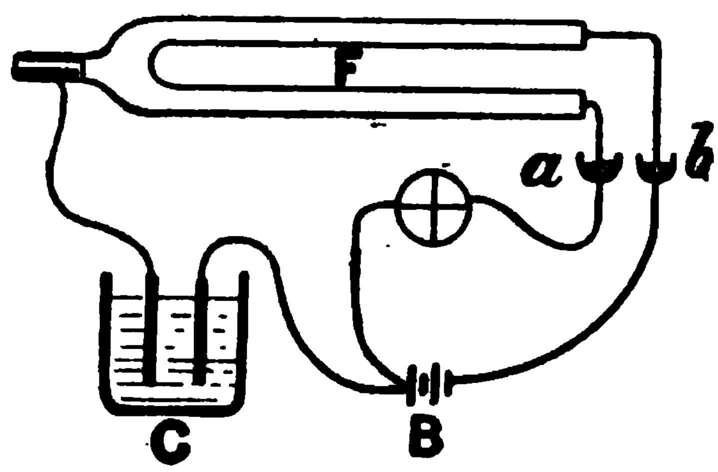


FIG. 518.

As the E.M.F. of the battery is increased it is found that at first the polarisation E.M.F. increases, and is almost exactly equal to the E.M.F. of the battery. As the primary E.M.F., however, gets larger the polarisation E.M.F. increases more slowly, but recent experiments seem to show that it goes on increasing very slightly even after the primary E.M.F. has reached such a value that an appreciable current will pass through the cell, so that no definite maximum of polarisation can be said to exist.

A phenomenon which is intimately related with that of polarisation, and which has already been referred to, is the minimum E.M.F. required to produce continuous decomposition. This E.M.F. must of course be greater than the polarisation E.M.F. for the corresponding current, for otherwise no current would pass through the cell, and hence no continuous decomposition would take place.

Numerous experiments on the minimum E.M.F. required to produce continuous decomposition have been made by Le Blanc, and some of the values he has obtained for solutions containing 1 gram equivalent per litre are given in the following table :—

Solution of	Minimum E.M.F. for Continuous Decomposition.	Solution of	Minimum E.M.F. for Continuous Decomposition (Unplatinised Electrodes).
$ZnSO_4$ . .	2.35 volts.	H_2SO_4 . .	1.67
$ZnBr_2$ . .	1.80 „	HNO_3 . .	1.69
$Pb(NO_3)_2$.	1.52 „	H_3PO_4 . .	1.70
$AgNO_3$ . .	0.70 „	HCl . . .	1.31
$Cd(NO_3)_2$.	1.98 „		
$CdSO_4$ . .	2.03 „		
$CdCl_2$ . .	1.78 „		

It will be noticed that, in the case of acids which on electrolysis evolve hydrogen at the kathode and, owing to secondary reactions, oxygen at the anode, decomposition starts with a potential difference of 1.7 volts. Experiments on the influence of the concentration of the solutions on the minimum E.M.F. required to produce continuous decomposition have shown that, in the case of solutions of acids where the value is about 1.7, this value is practically independent of the concentration. In the case of such acids as hydrochloric acid, however, where the value is considerably below 1.7, the minimum E.M.F. increases as the dilution is increased, and approaches the value 1.7 for very great dilutions. It is of interest to note that in a very dilute solution of HCl the gas liberated at the anode is no longer chlorine, but that a secondary reaction takes place and oxygen is evolved; so that at these great dilutions the electrolysis of HCl, as of the other acids, involves the evolution of H and O, and under these circumstances the minimum E.M.F. is the same for all.

CHAPTER XVI

CONTACT E.M.F. AND THE VOLTAIC CELL

545. Contact Electrification. — If a metal needle, A (Fig. 519), having the shape of half a quadrant electrometer needle, is suspended by a fine wire so as to be able to turn about a vertical axis through B, just above two metal semicircles, one of which, Z, is of zinc, and the other, C, of copper, then on electrifying the needle, if it is symmetrically arranged, no deflection will occur if the zinc and copper are insulated the one from the other. If, however, the zinc and copper are put in contact, either directly or through a conducting wire, the needle will be deflected. If the needle is charged with positive electricity, the deflection will be away from the zinc and towards the copper, thus indicating that the zinc is at a higher potential than the copper. This difference of potential between the zinc and copper, as indicated by the charged needle suspended over the metals, is said to be due to contact electrification. The magnitude of the contact difference of potential does not depend on the time the metals are in contact, nor on the area of the surface of contact; it does however depend on the nature of the metals, both chemical and physical, and on the temperature. The nature of the surfaces of the metals which are exposed to the air also has an important bearing on the magnitude of the contact difference of potential. A list of the metals can be drawn up such that any metal in the list when put in contact with any of the following metals is at the higher potential, but is at the lower potential when put in contact with any of the metals before it in the list. The following is such a list: Zinc, lead, tin, iron, copper, silver, gold. This list, which was first given by Volta, who discovered the contact effect, is called Volta's series.

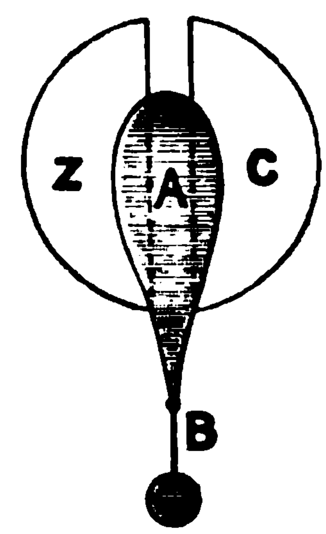


FIG. 519.

If three metals, A, B, and C, are put into contact in pairs, the difference in potential between any two is equal to the algebraic sum of the difference in potential produced by the contact of each of the metals with the third. Thus suppose the difference in potential produced by the contact of A and B is p_1, while that between B and C is p_2, then the difference of potential produced by the contact of A and C is p_1+p_2.

This law can be very clearly exhibited by means of diagrams in which the potential of a metal is represented by the height of a rectangle. Thus in the case of the three metals, tin, copper, and iron, the difference in potential between tin and copper is 0.5 volts, the tin being at the higher potential; hence, if we take 1 cm. to represent a volt, we draw the rectangles (*a*), Fig. 520, such that the height of the tin rectangle is 0.5 cm. greater than that of the copper rectangle. The difference in potential between tin and iron is 0.3 volts, so that, the tin rectangle being drawn the same height as before, the iron rectangle will be 0.3 cm. lower, as shown at (*b*). The difference in potential between the copper and the iron will be two-tenths of a volt, and if the rectangle for the copper is drawn of the same height as in (*a*), the rectangle representing the iron will be 0.2 cm. higher, that is, it will be of the same height as in (*b*).

If we imagine the copper and the iron both put into contact with the same piece of tin, then it is at once evident, from a consideration of Fig. 520 (*d*), that the difference in potential between the copper and the iron is the same as it is when they are put in direct contact. Thus

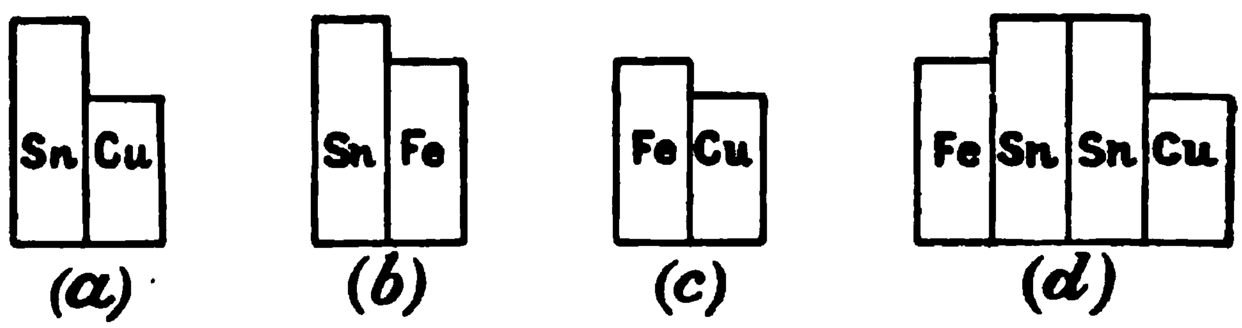


FIG. 520.

the difference in the potential of any two metals is the same, whether they are put in direct contact or whether they are joined by means of a wire composed of another metal.

It follows from the above law that if we arrange a circuit of which the parts are of different metals, but the first and last metals are the same, then there will be no difference in potential between these end portions.

If, however, the first and last metals are different, say A and B, the difference in potential between these metals being p, then the difference in potential between the end metals will be equal to p, although they are connected together by other metals.

It might at first sight appear, since we have two metals A and B at a difference of potential p, and that owing to the contact difference of potential they are kept at this constant difference, that on connecting A and B by means of a wire, a current would be set up in this wire. This, however, is not the case, for suppose we attempt to connect A and B by a wire of the metal A, then the difference of potential between the end of this wire and the metal B is p, but when the wire touches B,

owing to the contact, a difference in potential of p will be developed at the point of contact, and this difference of potential will prevent the difference of potential which exists between the metals A and B, forming the end of the chain, forcing electricity through the wire. The same can be shown to be true whatever the nature of the wire by which A and B are joined, so that by no arrangement of metals, all at the same temperature, can we obtain a current in a circuit which is composed exclusively of metals.

The case when we are dealing with the contact differences of potential between liquids, or between metals and liquids, is however quite different. Thus when copper is in contact with a solution of sulphuric acid, the copper is at the higher potential, while zinc, which is at the higher potential when in contact with copper, ought, if the liquid behaved as a metal would, according to the above law, to be at a higher potential than the sulphuric acid solution, instead of which it is at a lower potential. Hence it is possible to arrange a circuit composed partly of solid and partly of liquid conductors, such that a difference of potential exists between two parts of the circuit, even when these parts are connected by a conducting wire. Thus suppose we have a circuit composed of a plate of copper dipping in a solution of sulphuric acid, a plate of zinc also dipping in this solution, and a copper wire touching the zinc. The diagram of the potentials is shown in Fig. 521. The copper, Cu, is at a higher potential than the solution, H_2SO_4, while the solution is at a higher potential than the zinc, Zn. The zinc is at a higher potential than the copper wire, Cu′, so that the wire is at a lower potential than the copper plate. Hence by this arrangement we have got two portions of the same metal (copper), which, owing to contact differences of potential, are at different potentials, and, since when the copper wire is put in contact with the copper plate we are dealing with the contact of the same metal, and therefore no contact difference of potential is produced which would annul the tendency of the existing difference of potential to cause electricity to move in the circuit, we have here an arrangement suitable for producing an electric current.

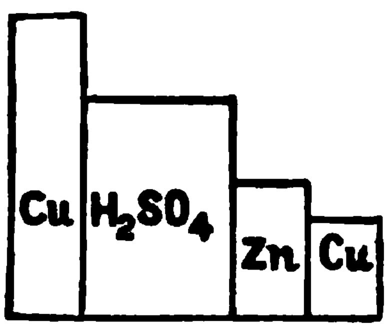


FIG. 521.

It is not even necessary that the contact of dissimilar metals occurs in the circuit, or even that two metals be employed, for a galvanic element can be produced in which no such contact of dissimilar metals occurs, or in which only a single metal is employed. Thus, when immersed in dilute sulphuric acid, copper is at a higher potential than lead, while when immersed in a solution of sodium sulphide (Na_2S), copper is at a lower potential than the lead. Hence if we have two glass vessels, one containing dilute acid and the other a solution of sodium sulphide, and place a strip of lead so that one end dips in the acid and the other end

dips in the solution of the sulphide, while one plate of copper is placed in the acid and another in the sulphide solution, then, as shown by the diagram in Fig. 522, the copper is at a higher potential than the lead in the acid in the vessel A, while the copper in the vessel, B, containing the sodium sulphide solution is at a lower potential than the lead. Hence the copper plate C is at a higher potential than the copper plate D, and if they are joined by a copper wire, a current of electricity will flow through the wire, although there is no contact of dissimilar metals.

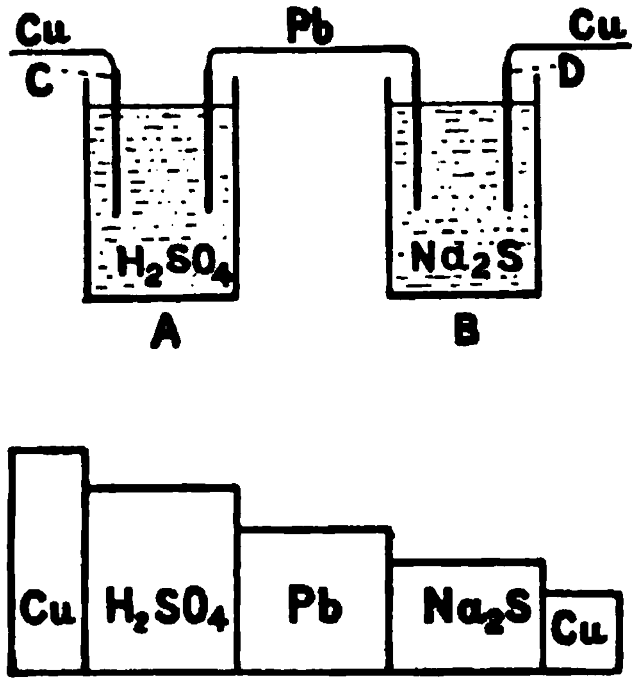


FIG. 522.

546. Magnitude of the Contact Difference of Potential.—In a voltaic cell consisting, say, of a plate of zinc and a plate of copper in dilute sulphuric acid solution, there are three different contacts between dissimilar materials, namely zinc/copper, copper/acid, and acid/zinc, and hence we have to deal with three contact differences of potential. The question as to the relative magnitude of these three contact differences is one which has occasioned an immense amount of discussion. The question does not lend itself to experimental decision, for no method has as yet been devised, which is free from all objection, for measuring the contact difference of potential between two bodies without the intervention of one or more other media, although, as we shall see in § 549, this can be got over if we accept the ionic theory. Thus, in the experiment of the charged electrometer-needle suspended over the zinc and copper quadrants described in § 545, what is measured is not the potential difference between the *zinc* and the *copper* but the difference in potential between the *air* in the neighbourhood of the zinc and that of the *air* in the neighbourhood of the copper. Hence if we indicate the true contact difference of potential between zinc and copper by Zn/Cu and so on, then what is actually measured is the sum of the three contact differences of potential, air/Zn + Zn/Cu + Cu/air. Thus if the two quantities air/Zn and Cu/air are not both zero or equal and opposite, it does not follow that the quantity Zn/Cu is not zero, or at any rate very small, so that the two metal quadrants may be really at the same potential, and we need not necessarily have two parts of a conductor at different potentials when the electricity is at rest.

Although it is impossible to obtain an experiment showing that the difference of potential of about 0.7 volts, which is observed between the air in the neighbourhood of a piece of zinc and a piece of copper which are in contact, is really due to the fact that the metals themselves are at

this difference of potential, yet the following considerations indicate that such a difference almost certainly cannot exist. Suppose that a copper and a zinc wire when joined together were actually at a difference of potential of 0.7 volts, and that a current C were passed through these wires in the direction from the copper to the zinc. The fall of potential along the combined wire can then be shown diagrammatically as in Fig. 523. The current passing in the direction OPQ, there is a continuous fall of potential along the copper wire OP according to Ohm's law, which is represented by the line AB. At P, however, where the copper and zinc meet, there will be a contact difference of potential which will raise the potential of the end of the zinc wire 0.7 volts above that of the copper, so that while the potential of the point P on the copper is represented by the point B on the curve, that of the point P on the zinc wire will be represented by C. In the zinc wire, from P to Q, the fall of potential will be regular and according to Ohm's law. Now as the current passes from the copper to the zinc it has to move *against* an E.M.F. represented by CB, that is, 0.7 volts, and the result is that an amount of work must be done represented by 0.7 C. joules in each second, for the electricity has at the junction been raised to a higher "level," and so its potential energy is increased. This increase in the electrical energy will take place at the expense of the heat of the junction, so that there will be an absorption of heat at the junction equal in electrical units to 0.7 C. watts. Or since a watt is equal to 0.2387 calories per second, there will be, if the current is one ampere, 0.167 calories absorbed per second. This absorption of heat will be reversible, so that on reversing the direction of the current the same quantity of heat would appear at the junction. Now in § 500 we have in fact considered this very problem when dealing with the Peltier effect, and it was there mentioned that the quantity of heat absorbed under the circumstances we have been considering is really only 1.6×10^{-4} calories per second. We are therefore led to the conclusion that no such difference of potential as 0.7 volts can really exist between a piece of copper and a piece of zinc when in contact, and that the difference of potential which has been called the contact difference of potential is mainly, at any rate, a difference of potential between the air in the neighbourhood of the two metals.

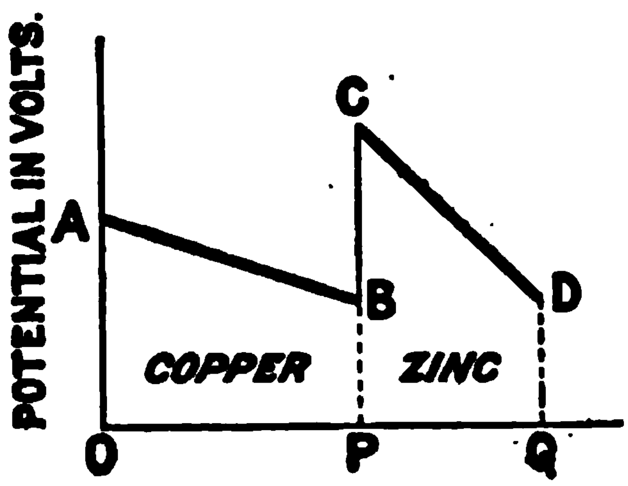


FIG. 523.

If we conclude that the actual difference of potential between two metals in contact is that which is deducible from the value of the Peltier effect, that is, of the order of 0.0007 volts, then, in most of the considerations as to the electromotive forces in circuits containing metals and

electrolytes, the contact difference of potential between metals may be neglected. We shall now proceed to consider more in detail the question of the contact difference of potential between metals and electrolytes, and between two electrolytes.

547. Electrolytic Solution Pressure.—The existence of a contact difference of potential between a metal and an electrolyte has been explained by Nernst on the ionic hypothesis in the following manner. Suppose a plate of a metal, say zinc, is dipped into a dilute solution of sulphuric acid, then Nernst supposes that there exists a tendency for the zinc in contact with the solution to enter into the ionic condition. Now each zinc ion must carry a charge of $+2\epsilon$, and as whenever a positive charge is produced we always find an equal and opposite negative charge formed at the same time, for each zinc ion that goes into solution in the water a negative charge -2ϵ will be developed on the plate; so that the zinc will become negatively electrified and the water positively, due to the presence of the positively charged zinc ions. If we further suppose that the tendency of the zinc to form ions is not indefinitely great, that is, that for a given metal and liquid there exists a definite pressure tending to make the metal assume the ionic condition, then the ions will be produced till the positive charge of the water and the negative charge on the plate exert such an electro-static force on each positively charged ion that is on the point of moving off into the water as to exactly balance the difference between the electrolytic solution pressure tending to cause the zinc to become ionised and the pressure of the zinc ions which have already gone into solution. This hypothesis of a solution pressure is analogous to the vapour pressure of a liquid, for evaporation goes on till the pressure of the vapour is equal and opposite to the pressure which tends to make the liquid particles assume the gaseous form. Since the charge on each ion is very great (§ 539), for 1/1000 of a milligram of zinc ions carry a charge of about 3 coulombs, it is not necessary to assume that a weighable quantity of the zinc passes into the water in the ionic condition to account for the observed difference of potential.

If the metal is in a solution which already contains its own ions, as for instance a plate of zinc in a solution of zinc sulphate, then the ions already present in the solution will offer an opposition to the entrance of any more of the same ions, or we may suppose that they exert a pressure tending to drive them out of the solution, so that the difference of potential necessary to produce equilibrium is less than when no zinc ions are in the solution. If the electrolyte is such that the tendency of the ions it already contains to leave the solution is greater than the solution pressure of the metal, then a certain number of the ions will leave the solution and become electrically neutral on the metal. In this way the metal will acquire a positive charge and the solution a negative one. As before, the positive charge on the plate will exert a repulsive force on the positively charged ions remaining in the solution, while the negatively charged solution will exert an

attraction, so that both of these forces will oppose the passage of the ions on to the metal, and a state of equilibrium will be set up in which the electro-static force and the solution pressure of the metal are together equal and opposite to the force with which the ions in the liquid tend to leave the solution and attach themselves to the metal.

Contact differences of potential not only exist between two different electrolytes in contact, but also between two solutions of the same electrolyte if the concentration is different. Nernst has explained the difference of potential existing between two solutions of the same salt, when the concentrations differ, in the following way. Suppose that a strong solution of, say, hydrochloric acid is in contact with pure water, then the acid will diffuse into the water. Since, on the ionic theory, the hydrogen ions and the chlorine ions are regarded as being capable of independent motion, and since their velocities of migration are regarded as different, that of the hydrogen being the greater, the H ions will travel faster into the water than the Cl ions. Hence, as the H ions carry a positive charge, the water will become positively charged, owing to the presence of an excess of H ions, and the solution negatively, owing to the excess of Cl ions. The process will not go on indefinitely, for as the water becomes positively charged an electro-static repulsion will be produced, tending to check the advent of the positively charged H ions, and to accelerate the negatively charged Cl ions. Thus the H ions will diffuse more quickly at first, till the difference of potential produced between the water and the solution is so great that the electro-static forces on the charged ions cause the H and Cl ions to diffuse at the same rate. As the diffusion continues, the number of ions in the weaker solution will increase, and hence the tendency of the ions from the stronger solution to move into the weaker solution will also decrease, and the difference of potential necessary to prevent the H ions diffusing more quickly than the Cl ions will be less. In other words, the contact difference of potential will decrease when, owing to diffusion, the concentrations of the two solutions become more nearly equal.

In the case of electrolytes it is impossible to make a list such as Volta's list for the metals, so that it is possible to arrange a series of electrolytes to form a circuit which can produce a current through the circuit. Such a circuit (Fig. 524) consists of a fairly strong solution of potassium chloride (KCl), a weak solution of KCl, a weak solution of hydrochloric acid (HCl), a fairly strong solution of HCl, this solution being in contact with the strong solution of KCl, when a current will circulate through the solutions as shown diagrammatically in Fig. 524.

Fig. 524.

A current-producing circuit cannot, however, be formed by a series of differently concentrated solutions of the same electrolyte.

548. The Capillary Electrometer.—The measurement of the contact difference of potential between a metal and a liquid has been rendered possible owing to a discovery of Lippmann's, who found that when a current is passed across the surface separating mercury and dilute sulphuric acid, in the direction from the acid to the mercury, the area of the surface of separation tends to decrease, as if the surface tension (§ 157) had increased. Conversely, if by any means the area of the surface of separation is varied, a current will be produced in a wire connecting the mercury and the solution. The effect of the application of an E.M.F across the surface of separation of mercury and dilute sulphuric acid on the surface tension has been applied by Lippmann to the construction of an electrometer. A simple form of capillary electrometer is shown in Fig. 525. A glass tube, AB, is drawn off at the bottom into a fine capillary, C. To the side of this tube is connected a reservoir, D, by means of a flexible indiarubber tube. AB and D are filled with mercury, so that by raising D the height to which the mercury rises in AB can be varied, and hence also the pressure which acts tending to drive the mercury out through the capillary. The end of the capillary dips in a vessel, E, containing a dilute solution of sulphuric acid above and some mercury below. Platinum wires are fused through the tube AB at X, and through the vessel E at Y, so that an E.M.F. may be applied between the mercury in AB and that in E. Since the angle of contact between mercury and sulphuric acid solution and glass is greater than 90°, the surface tension tends to drive the mercury up the capillary tube, and a condition of equilibrium will be reached when the capillary force is just balanced by the head of mercury in AB. The position of the meniscus separating the mercury and the solution in the capillary is observed by means of a microscope, T. Since, as we have seen in § 160, the pressure which the capillary forces will sustain depends on the diameter of the capillary, and as the capillary is never cylindrical, but is slightly conical, the meniscus will move till it reaches a point in the capillary where the diameter of the bore is such that the capillary forces just balance the head of mercury. If the head of mercury is increased the meniscus will move down into a narrower part of the capillary, and *vice versa*. If after equilibrium has been attained the surface tension is in any way reduced, the meniscus will move down to a narrower part of the capillary till the increase of the capillary forces due

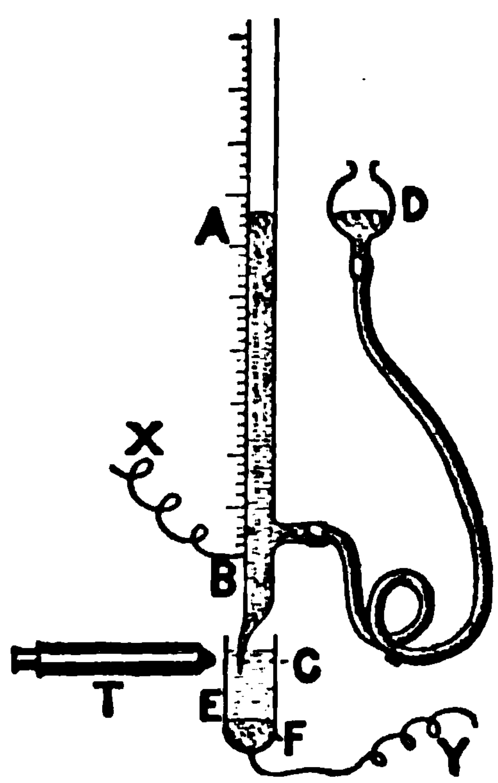


Fig. 525.

to the reduced diameter just makes up for the reduction in the surface tension, and to bring the meniscus back to its original position, as indicated by the cross wires of the microscope, the head of mercury will have to be reduced by lowering the reservoir, D; and the decrease in the head as measured by a scale fixed alongside AB will be a measure of the decrease in the surface tension. In the same way, if the surface tension increases the meniscus will move up to a wider part of the capillary, and to bring it back to the fixed position the head of mercury will have to be increased. If the head of mercury is so adjusted that the meniscus is in the sighted position when the wires X and Y are connected together, so that no external E.M.F. is acting across the meniscus, it is found, on applying such an E.M.F., that the meniscus moves, and that to bring it back to its sighted position the head of mercury in AB has to be altered. For small differences of potential the alteration which has to be made in the pressure applied to the mercury to bring the meniscus back to its sighted position is proportional to the applied E.M.F. The working of the capillary electrometer has been explained by von Helmholtz in the following manner. At the surface of separation between the mercury and the solution there will exist a contact difference of potential, and since both the mercury and the solution are conductors, and at their surface of separation they are very near together, the difference of potential will cause an accumulation of electricity on the two sides of the bounding surface. For instance, if we have two plates of conducting material separated by a thin sheet of a dielectric, say ebonite, then if a difference of potential is produced between the conductors we know that the arrangement will act as a condenser, and the charge on the two conductors will accumulate on the surfaces which are in contact with the ebonite. The same thing occurs in the case of the mercury and the solution. Since a difference of potential is maintained, how it is maintained is immaterial, and we do not know for certain, although in § 547 an explanation has been given, and the effect is just as if a non-conducting sheet had been interposed between the mercury and the solution, and then the two had been charged to a difference of potential equal to the contact difference of potential.

Since the mercury is positive to the solution, we must suppose that this double layer of electricity on the surface of separation consists of positive electrification on the mercury side, and negative on the solution side.

We have now to consider the influence of this double layer on the surface tension of the surface of separation. In § 157 it was shown that if T is the observed surface tension of the surface separating two media, and the area of this surface is increased by an amount s, the work which has to be done is sT. Now the surface of separation between the mercury and acid solution with its double layer may be regarded as a condenser, of which the two armatures are charged to a potential difference e, where

e is the contact difference of potential between the mercury and the solution. Now, in any condenser of which the plates are kept at a constant difference of potential, the electrical forces tend to *increase* the capacity of the condenser. Hence, in the case of the double layer, there is a tendency for the capacity to increase, that is, for the area of the double layer to increase. The result is that, on account of the electrical forces, the area of the surface of separation between the mercury and the solution tends to increase, and so the electrical forces will reduce the amount of work which has to be done against the purely surface tension effects when the area of the surface of separation is increased. Thus if T' is the value the surface tension would have, suppose no electrical double layer were present, the work done in increasing the area of the surface of separation by an amount s would be sT'. Therefore sT, the actual amount of work done in increasing the surface, is less than sT', the amount of work which would have to be done if no electrical double layer existed, by the amount of work done by the electrical forces owing to the increase in the capacity of the double layer. Thus T, the actual surface tension, is less than T', the surface tension suppose no double layer existed, the presence of the double layer causing a diminution in the observed surface tension.

Suppose the contact difference of potential between the mercury and the solution be e, the mercury being at the higher potential. Then, if an external E.M.F. be applied, so that the wire X is positive, the difference of potential between the mercury and solution will be greater than e by the amount of the applied E.M.F., and hence the charges on the double layer will be increased, so that the surface tension will be decreased, and to keep the meniscus in its sighted position the head of mercury in AB must be reduced. If, however, the applied E.M.F. is in such a direction that it acts in the opposite direction to the contact difference of potential at the meniscus, then the strength of the double layer will decrease, and hence the surface tension will increase. This increase will go on till the applied E.M.F. is exactly equal and opposite to the contact difference of potential, for when this occurs there will be no double layer, and hence the surface tension will possess the value which it would have if no electrical charges were present. If the applied E.M.F. is further increased, but still in the opposite direction to the contact difference of potential, then a double layer will again be formed, but with the negative charge on the mercury side. This inverted double layer will cause a decrease in the surface tension, since the presence of such a double layer must decrease the surface tension whichever side is positive. Hence, by applying an external E.M.F. so as to make the mercury negative, and increasing it till the surface tension, as indicated by the pressure which has to be applied to bring the meniscus to its sighted position, is a maximum, we shall have that the applied E.M.F., when this maximum is reached, will be exactly equal and opposite to the contact difference of potential between

the mercury and the sulphuric acid solution. In this way, Lippmann found that the contact difference of potential between mercury and sulphuric acid solution was about 1.0 volt.

The value of the surface tension when the applied E.M.F. is exactly equal and opposite to the contact difference of potential will be the value of the surface tension unaffected by electrical disturbances, and Ostwald has found that this maximum surface tension is frequently independent of the nature of the electrolyte, although the values obtained in the ordinary way differ considerably with different electrolytes.

549. Values of the Contact Differences of Potential between Metals and Liquids.—Having, by means of the capillary electrometer, determined the contact difference of potential between mercury and any given electrolyte (other electrolytes besides sulphuric acid may be employed), the value of the contact difference of potential between any other metal and the electrolyte can be obtained. For, suppose we have a metal M, and we require to find the contact difference of potential between this metal and sulphuric acid. If a plate of the metal M is dipped in a vessel containing a solution of sulphuric acid and also some mercury, and if e_1 is the contact difference of potential between mercury and the solution, and e_2 that between M and the solution, the difference of potential between the mercury and the metal will be equal to $e_1 - e_2$. Hence, since we can measure this difference of potential between the metal and the mercury,[1] and we know e_1, we can calculate e_2. The following values for the contact differences of potential have been obtained in this way, the potential of the electrolyte being taken in all cases as zero, and the solution containing one gram equivalent per litre :—

CONTACT DIFFERENCES OF POTENTIAL.

Metal.	Electrolyte.				
	H_2SO_4.	Sulphate.	HCl.	Chloride.	Nitrate.
Zinc . . .	−0.62	−0.52	−0.54	−0.50	−0.47
Cadmium . .	−0.22	−0.16	−0.24	−0.17	−0.12
Copper . .	+0.46	+0.52	+0.35	...	+0.62
Silver . .	+0.73	+0.97	+0.57	...	+1.06
Mercury . .	+0.86	+0.98	+0.57	...	+1.03

550. The Voltaic Cell.—If a plate of copper and one of zinc are dipped in a vessel containing a dilute solution of sulphuric acid, then, from the table given above, it will be seen that if the potential of the solution is taken as zero, the potential of the copper will be +0.46

[1] Supposing that the contact difference of potential between the metal M and mercury is so small as to be negligable.

volts, and that of the zinc will be −0.62 volts. Hence the copper will be at a potential of 1.08 volts higher than that of the zinc, so that if the copper and zinc plates are joined by a conducting wire, this wire will be traversed by an electric current. This arrangement is called a simple galvanic or voltaic cell or battery.

When the copper plate is connected to the zinc by the conducting wire, positive electricity will flow from the copper to the zinc, and so the potential of the copper with reference to the solution will decrease. The result will be that the electro-static force which opposed the precipitation of the hydrogen ions of the electrolyte on the copper electrode will be reduced, and so more of these ions will be able to discharge themselves on the copper. In the same way, the increase of the potential of the zinc due to the passage of electricity round the conducting wire will decrease the electro-static force which opposes the solution pressure of the zinc, and so more zinc ions will be able to go into solution, carrying their positive charge with them. Thus the passage of a current in the wire from the copper pole to the zinc pole is accompanied by the solution of the zinc at one electrode and the liberation of hydrogen at the other. For every equivalent of the zinc dissolved an equivalent of hydrogen will be liberated, and the passage of 96,550 coulombs of electricity in the wire will be accompanied by the solution of the electro-chemical equivalent of zinc, and the liberation of the electro-chemical equivalent of hydrogen. Thus the cell behaves, as far as the chemical changes which take place within it are concerned, exactly as if it were an electrolytic cell through which a current is sent by some external agency.

The potentials of the different portions of the simple cell, considered before the external circuit was closed, may be represented by the curve (*a*), Fig. 526, in which the ordinates represent the potentials of the different portions of the circuit. It will be seen that the copper, the zinc, and the solution are each at the same potential throughout. When the external circuit is closed, so that a current flows in the external wire, the end A of the copper wire being in contact with the end F of the zinc, these two points will be at the same potential. We are here neglecting the contact difference of potential between the copper and the zinc, since this is so very small compared with the contact differences between the metals and the solution. Since there is a current flowing in the copper wire in the direction BA, there will, by Ohm's law, be a fall of potential along the wire as indicated by the line AB, Fig. 526 (*b*). In the same way, as a current is passing through the zinc in the direction FE, there will be a fall of potential along this wire as shown by FE. Similarly, since Ohm's law applies to the electrolyte, there will be a fall of potential in the liquid between D and C. Study of the figure will show that, although the copper plate is still at a potential of 0.46 volts above that of the solution in its immediate neighbourhood, as represented by BC, and that of the zinc plate is at

0.62 volts below that of the electrolyte in its immediate neighbourhood, as represented by DE, yet, owing to the fall of potential along the electrolyte so that the end D is at a higher potential than the end C, the difference of potential between the copper and zinc plates, that is, between the poles of the cell, is less when a current is passing than when the cell is on open circuit. The change produced on this account can immediately be calculated, for if r is the resistance of the electrolyte between the copper and zinc plates, that is, the resistance of the cell, then, when a current C is passing, the fall of potential will by Ohm's law be equal to rC. Hence the potential between the poles of the cell, when it is producing a current C, will be $1.08 - rC$.

The above discussion, in the case of such a cell as the one described, only applies to the first moment of closing the circuit, for after an appreciable current has passed, polarisation effects will occur which will decrease the available E.M.F. The polarisation with which we are

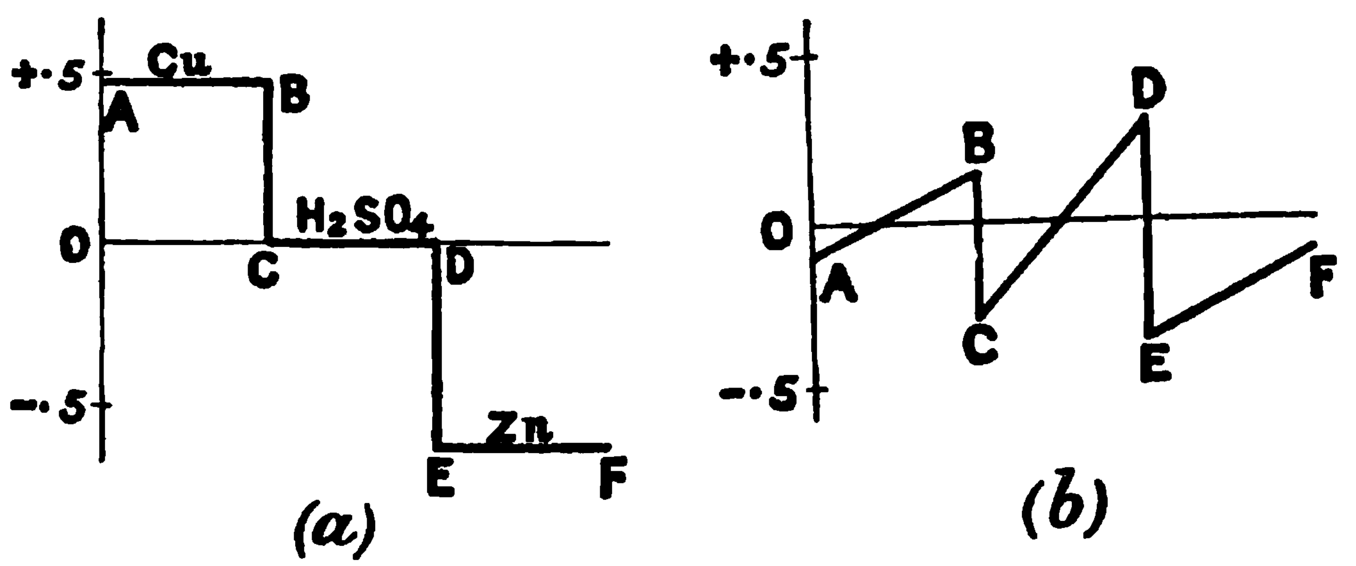


FIG. 526.

here concerned is of the second kind considered in § 544, that is, it is due to the effect of the hydrogen liberated at the copper plate and to a less extent due to the dissolved zinc ions not diffusing away from the zinc, so that the effect of the solution pressure of the zinc will be diminished owing to the presence of an increased number of zinc ions in the solution near the plate. The polarisation due to the hydrogen is produced, in the first place, by the copper becoming coated with a layer of fine gas-bubbles which increases the resistance of the cell, owing to the diminished surface of the copper available for the passage of the current; and, in the second place, by the copper becoming coated with hydrogen, the positive plate becomes practically a plate of hydrogen, and, since hydrogen has a smaller contact difference of potential with the solution than copper, the E.M.F. of the cell is decreased. A cell such as the above, in which, owing to polarisation, the E.M.F. decreases rapidly when a current is allowed to pass, is called an inconstant cell. In order to obviate the polarisation, we must so choose our electrolyte that

the chemical processes which take place when a current passes do not cause the accumulation of the ions on the electrodes in such a way as to increase the resistance or decrease the contact differences of potential between the electrodes and the electrolyte. Cells which fulfil this condition more or less completely and are, at any rate at first, free from polarisation, and of constant E.M.F. even after sending a current, are called constant cells, and we shall now proceed to describe some of the commoner forms of such cells. It must be remembered that there must always be a decrease in the difference of potential between the poles numerically equal to rC, where r is the resistance of the cell and C the current which is passing. This apparent reduction in the E.M.F. is due to the fact that the liquid of the cell has appreciable resistance and hence, by Ohm's law, a certain proportion of the available E.M.F. has to be employed in driving the current through the liquid, that is, in moving the H ions, in the simple cell previously considered, to the copper pole, and the SO_4 ions to the zinc pole. This effect being quite independent of the polarisation, will not be effected by any change calculated to diminish the polarisation, and can only be reduced to a minimum by making the resistance of the liquid, that is, r, as small as possible.

551. The Daniell Cell.—This form of cell consists of a zinc plate dipping in a solution of sulphuric acid or zinc sulphate and a copper plate in a solution of copper sulphate, the two solutions being prevented from mixing by the interposition of a partition composed of porous earthenware. In some forms of the cell, called gravity Daniells, the porous partition is done away with, the copper plate Cu (Fig. 527) being placed at the bottom of a glass vessel and covered with a saturated solution of copper sulphate. The zinc sulphate, which has a less density than the copper sulphate solution, floats on the top of the latter and, since convention currents cannot be formed and the process of diffusion is very slow, the solutions do not mix for some time. The negative plate is formed by a horizontal disc of zinc, Zn. The connection with the copper plate is made by means of a wire, A, which passes down to the copper and is enclosed in an insulating tube, generally of glass.

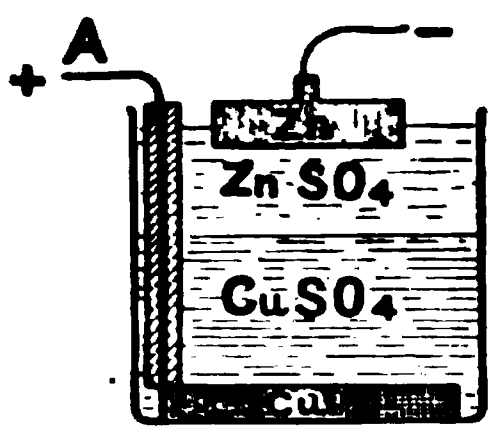


FIG. 527.

The E.M.F. of the Daniell cell is due to the sum of the contact differences of potential, Cu $CuSO_4$, $CuSO_4$ $ZnSO_4$, $ZnSO_4$/Zn, Zn Cu, and is about 1.096 volts. Since it is very probable that the contact differences of potential Zn/Cu and $CuSO_4$/$ZnSO_4$ are very small, we may calculate the E.M.F. of the cell by means of the table given on page 815. Thus Cu $CuSO_4 = +0.52$ and $ZnSO_4$/Zn $= +0.52$, so that the E.M.F. is 1.04 volts, a number agreeing approximately with the value

1.096 obtained by direct measurement. The E.M.F. of a Daniell cell is increased by increasing the strength of the copper sulphate solution and by diluting the zinc sulphate solution.

When the external circuit of a Daniell cell is closed, so that a current passes, the zinc goes into solution as zinc sulphate, while the kation of the copper sulphate solution, that is, the copper, is deposited as metallic copper on the copper plate of the cell. It will thus be seen, since the deposition of copper on the copper kathode will in no way affect either the resistance of the cell or the contact difference of potential between the copper sulphate solution and the copper, that the E.M.F. of this form of cell will not be decreased on account of polarisation. When the cell sends a current the SO_4 ions, each carrying a charge of -2ϵ, move from the neighbourhood of the copper pole to the zinc pole, and the copper ions which are left at the copper pole are deposited, giving up their positive charge. At the same time the zinc ions enter the solution from the zinc pole, each carrying a positive charge, and these positive ions, together with the negative SO_4 ions which have migrated from near the copper pole, being in equivalent proportions in the solution, prevent the solution becoming charged.

552. The Grove Cell.—The positive pole of this cell consists of a plate of platinum in a strong solution of nitric acid, and the negative pole is a zinc plate in a fairly strong solution of sulphuric acid (1 of acid to 10 of water), the liquids being separated by a porous earthenware partition. The E.M.F. of this cell is about 1.97 volts. When a current passes, the zinc goes into solution, forming zinc sulphate with the SO_4 ions of the sulphuric acid solution; while the H ions migrate, each carrying its positive charge, to the platinum plate, where they give up their charge and thus transport the current through the cell. The hydrogen is not, however, given off at the platinum, but a secondary reaction takes place between it and the nitric acid, which results in the combination of the hydrogen with part of the oxygen of the acid to form water, and leaves an oxide of nitrogen in the solution. The E.M.F. of the cell gradually falls off, owing to the exhaustion of the nitric acid allowing polarisation to take place, as well as the gradually increasing concentration of the zinc ions in the solution diminishing the potential fall from the acid to the zinc.

The Bunsen cell is the same as the Grove cell, except that the positive pole consists of a plate of gas carbon. A solution of chromic acid is sometimes used in place of the nitric acid, the action being of a similar nature. Since the presence of chromic acid near the zinc does not materially alter the solution pressure of the zinc ions, and does not produce any secondary chemical action with the zinc, the porous cell separating the chromic acid and the sulphuric acid solution may be omitted. This form of cell is called the chromic acid cell or, since bichromate of potash is sometimes used in place of the chromic acid, the bichromate cell.

553. The Leclanché Cell.—The positive pole of this cell consists of a plate of carbon packed round with a mixture of powdered carbon and manganese dioxide (MnO_2). The negative pole consists of zinc, and the electrolyte is a solution of ammonium chloride or sal-ammoniac (NH_4Cl). During the passage of the current Cl ions move towards the zinc, forming zinc chloride, while the NH_4 ions move towards the carbon kathode, where they break up into ammonia (NH_3), which dissolves in the solution, and hydrogen, which combines with part of the oxygen of the MnO_2 to form water and MnO. The E.M F. of this cell on open circuit is about 1.6 volts, and with any but small currents the cell polarises rather rapidly. After a short rest, however, the cell recovers, and it possesses the great advantage that no chemical action goes on when no current is passing, and as there is only one kind of electrolyte, diffusion of one electrolyte into the other, which always occurs in time when two electrolytes are employed, does not occur.

554. The Clark Cell.—This cell is not used for sending any but the very smallest currents, but is employed as a standard of E.M.F., and consists of an amalgamated zinc anode in a saturated solution of zinc sulphate and a mercury kathode covered with a paste formed by mixing mercurous sulphate with saturated zinc sulphate solution.

The form of the Clark cell recommended by the British Board of Trade is shown in Fig. 528. The mercury is placed at the bottom of a small glass tube, contact being made by means of a platinum wire, which either passes down a glass tube or is fused through the bottom of the glass. The mercurous sulphate paste forms a layer, B, on the surface of the mercury, while the saturated $ZnSO_4$ solution C floats on the top of this paste. The zinc rod passes through a disc of cork, D, and the remainder of the tube above the cork is filled with marine glue, which serves to seal the cell and thus prevent the evaporation of the solution. Since the constancy of the E.M.F. of the cell depends on the zinc sulphate solution remaining saturated, even when the temperature rises so that the solubility of the salt increases, it is usual to pack the space above the paste with small crystals of zinc sulphate.

FIG. 528.

The E.M.F. of this form of cell, at a temperature of 0° C., is 1.4488 volts, while the E.M.F. at a temperature $t°$ C. is given by the expression

$$E_t = 1.4322 - 116 \times 10^{-5}(t-15) - 10^{-6}(t-15)^2.$$

The somewhat large change in the E.M.F. with temperature is no doubt partly due to the change in the solubility of the zinc sulphate with temperature, and since when the temperature rises, in order that the formula given above may hold, it is necessary that the solution should remain

saturated at this new temperature by some of the crystals dissolving, it is important that this solution should be assisted in every way, and that the zinc anode at any rate should always be in a portion of the solution which is saturated. In the form of cell shown in Fig. 528 there is a tendency for the denser saturated solution to accumulate on the surface of the paste when the temperature is raised, so that the solution surrounding the zinc may not be saturated. With a view of remedying this defect, the form of cell shown in Fig. 529 has been devised. The cell is contained in an H-shaped glass tube, the mercury A being placed at the bottom of one of the limbs. An amalgam containing about 10 per cent. of zinc is used for the anode, and is placed at the bottom of the other limb, Zn. The use of an amalgam in the place of pure zinc does not affect the E.M.F. of the cell, for it has been found that when an alloy of two metals, such as zinc and mercury, is placed in contact with an electrolyte it takes up a difference of potential from the liquid which corresponds to that which would occur if the metal of which the solution pressure is the greater (Zn) were alone present, so long as the other metal is not in great excess. Communication is made with the mercury and the amalgam by means of platinum wires, which are fused through the bottoms of the limbs of the glass containing vessel. The mercurous sulphate paste B is placed on the top of the mercury, and a layer of zinc sulphate crystals, C, is placed above the amalgam, so that when the temperature alters the solution may remain saturated. In this form of cell it will be noticed that, as the zinc is below the solution, the denser saturated solution will fall and cover the zinc, so that the solution near the zinc surface will always be saturated.

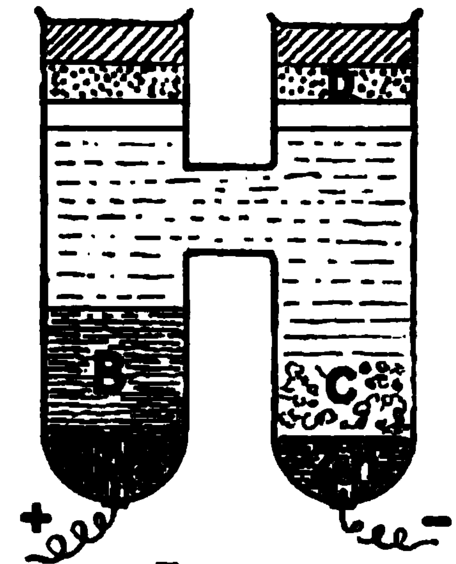


FIG. 529.

555. The Cadmium Cell.—This is another form of cell used as a standard of electromotive force, and not to furnish a current, and it possesses the advantage over the Clark cell that the E.M.F. changes much less with temperature. It consists of a mercury kathode covered with a paste formed of mercurous sulphate and a saturated solution of cadmium sulphate. The anode consists either of a rod of cadmium or of an amalgam of cadmium, while the electrolyte in which the cadmium is placed is a saturated solution of cadmium sulphate.

The arrangement of the materials to form the cell is exactly the same as that adopted in the case of the Clark cell. The E.M.F. of the cadmium cell, at a temperature t° C., is given by the expression

$$E_t = 1.0184 - 3.8 \times 10^{-5}(t - 20) - 0.065 \times 10^{-5}(t - 20)^2.$$

556. Reversibility of Cells.—If a current is passed through a Daniell cell by means of an external source, so as to enter the cell at the

copper pole and leave by the zinc pole, the chemical reactions which will take place will be exactly the reverse of those that occur when the cell itself sends a current, for the copper will be dissolved to form copper sulphate, and the zinc deposited from the zinc sulphate solution. Hence, passing an electric current in the reverse direction through a Daniell cell of the type Zn, $ZnSO_4$, $CuSO_4$, Cu, is accompanied by chemical changes such that, when the cell is itself allowed to send a current, the inverse chemical changes take place. Thus part at any rate of the energy spent in sending the original current through the cell is stored up in such a way that it may at a future time be reconverted into electrical energy. Any other form of cell may be used in the same way, provided the products of the chemical actions which take place during the working of the cell are retained either on the electrodes or in the electrolyte, and are not given off. The simple voltaic cell, consisting of a plate of copper and one of zinc in dilute sulphuric acid, is not reversible, since the hydrogen which is evolved at the kathode when the cell is sending a current escapes in the gaseous form. A form of cell which is specially designed to store up electrical energy, so that it can be recovered at a subsequent time in the form of a current, is called a storage cell.

557. The Storage Cell.—The commonest form of storage or secondary cell consists of two lead grids, the interstices being filled with lead sulphate formed by making a paste with one of the oxides of lead, litharge or red lead, and dilute sulphuric acid. These plates are immersed in a dilute solution of sulphuric acid, and then a current is passed through the cell from one plate to the other. During the passage of the current, the hydrogen ions of the sulphuric acid travel to the kathode, where they react on the lead sulphate, forming sulphuric acid and metallic lead, which remains in the interstices of the plate in a very spongy condition. The SO_4 ions travel to the anode, where they also react on the lead sulphate, forming peroxide of lead and sulphuric acid according to the equation

$$PbSO_4 + SO_4 + 2H_2O = PbO_2 + 2H_2SO_4.$$

The peroxide of lead is left in the interstices of the grid.

When nearly, if not quite all, the lead sulphate on the grids has been changed in this way, the hydrogen ions will be liberated at the kathode in the form of gas, while at the anode, owing to the secondary reaction between the SO_4 ions and the water of the solution, which has already been referred to when considering the electrolysis of dilute sulphuric acid between platinum electrodes, oxygen is liberated. When this evolution of gases occurs, the cell is no longer working in a reversible manner, and it has received the maximum charge of which it is capable.

If, after being charged in this way, the plates are connected by a conducting wire, a current will be obtained in the reverse direction to that employed to charge the cell, the chemical changes taking place in the

reverse direction, the spongy metallic lead becoming converted into the sulphate, and the peroxide also forming the same compound.

The E.M.F. of a freshly charged accumulator is about 2.1 volts, which gradually falls to about 1.8 volts as the discharge goes on. The lead accumulator is a wonderfully efficient means of storing energy, since about 80 per cent. of the energy spent in charging the cell is recoverable if the discharge takes place within a fairly short interval after the charge. The disadvantage of the lead storage cell lies in the fact that, owing to the considerable changes in volume which take place in the active material during charge and discharge, the grids disintegrate pretty rapidly, and hence the expense of the renewals of these plates has to be taken into account when considering the efficiency of the cells from a practical standpoint.

CHAPTER XVII

ENERGETICS OF THE VOLTAIC CELL

558*. Source of the Energy of the Current given by a Voltaic Cell.—We have considered the question as to how the E.M.F. of a voltaic cell is produced, and now we have to consider more in detail from whence the energy necessary for the maintenance of a current is derived. This energy is evidently, in part at any rate, derived from the energy of the chemical processes which take place in the cell during the time when it is sending a current. In § 228 we have considered the energy which is liberated or absorbed during certain chemical changes, and the question arises as to the connection between the total quantity of energy which is evolved as heat, when the reaction takes place without the production of an electric current, and the energy represented by the current when this is produced. It was thought for some time that the whole of the energy corresponding to any chemical change was converted into electrical energy when the change took place in a voltaic cell, and the fact that the E.M.F. of the Daniell cell, when calculated on this hypothesis from the thermo-chemical data for the chemical changes which take place in this cell, agreed very well with the value as obtained by direct measurement, supported this view.

Thus in § 228 we have seen that when 65 grams (one gram atom) of zinc are dissolved in dilute sulphuric acid according to the equation $Zn + H_2SO_4 = ZnSO_4 + H_2$, 38,066 calories are evolved. Experiment has shown that when 63 grams of copper are converted into copper sulphate in solution in water according to the equation $Cu + H_2SO_4 = CuSO_4 + H_2$, 12,500 calories are absorbed, so that 12,500 calories are evolved when $CuSO_4$ is split up into Cu and H_2SO_4. Hence when one equivalent, that is, since zinc is a diad, 65/2 grams of zinc are converted into the sulphate, while at the same time one equivalent of copper (63/2 grams) is deposited from the sulphate, $19033 + 6250 = 25283$ calories are on the whole evolved.

Now the reactions considered above are those which go on in the Daniell cell when it is sending a current, and we have seen that the quantities of zinc and copper considered above, are dissolved and precipitated respectively when 96,550 coulombs of electricity pass through the cell. If the E.M.F. of the cell is E volts, then the passage of 96,550 coulombs of electricity will correspond to $96550 \times 10^7 \times E$ ergs, for one

volt is equal to 10^8 *c.g.s.* units, and one coulomb is 10^{-1} *c.g.s.* units. If, then, the whole of the energy corresponding to the chemical reaction is converted into electrical, we shall have, since 25,283 calories is equal to $25283 \times 4.2 \times 10^7$ ergs, or 1068×10^9 ergs,

$$96550 \times 10^7 E = 1068 \times 10^9 \text{ ergs}$$
$$E = \frac{106800}{96550}$$
$$= 1.106 \text{ volts.}$$

Now direct measurement has given the value 1.096 volts for the E.M.F. of a Daniell cell, so that in this case it would seem that the electrical energy of the cell is equal to the chemical energy corresponding to the reactions which go on in the cell during the passage of the current.

When, however, the same method of calculation came to be applied to other forms of cells it was found that the E.M.F.'s calculated on this hypothesis differed from the observed values by more than could be accounted for by errors of experiment. The reason for these differences was shown by Helmholtz to be due to the fact that the hypothesis that the electrical and chemical energies were in all cases exactly equal was not true. He showed that this was only true in the case of cells in which the E.M.F. does not vary with the temperature, the Daniell being a cell of this kind.

In order to see the reason for this, we may consider the case of a reversible cell in which all the chemical changes that take place when the cell is allowed to send a current can be reversed when a current is sent through the cell in the reverse direction.

Suppose that when the temperature of the cell is T_1 (on the absolute scale) the E.M.F. is E_1, and that when the temperature is reduced to T_2 the E.M.F. falls to E_2. If now, when the temperature of the cell is T_1, it be allowed to send a current till Q units of electricity have passed, the work done is QE_1, and is represented by the area of the rectangle ABMO (Fig. 530). Now let the temperature of the cell be reduced to T_2, so that the E.M.F. is E_2, and let a quantity of electricity Q be passed through the cell in the reverse direction. The work which will have to be done will be QE_2, and is represented by the rectangle CDOM. During the passage of this electricity in the reverse direction, the chemical changes which took place in the cell during the time when it was sending a current will be exactly reversed, so that if the cell be now heated up to the temperature T_1 it will be in exactly the same condition as that in which it was at the start, that is, it will have been carried through a cycle of operations (§ 260).

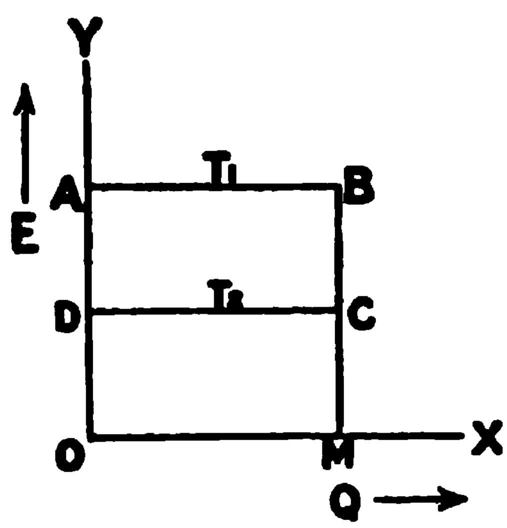


FIG. 530.

Since, as the cell is a reversible cell, none of the products of the chemical changes which go on during the passage of a current escape from the cell, and also since the elements carry their atomic heats into their compounds, it will require the same quantity of heat to raise the temperature of the cell from T_2 to T_1 as it did to cool it from T_1 to T_2, and hence these operations exactly balance one another and need not be considered. Also any external work done on account of change of volume will be negligable. A consideration of the diagram shows that an amount of external work represented by the rectangle ABCD has been done during the cycle, and since the chemical state of the cell is the same at the end as at the start, it follows that this work cannot have been done at the expense of chemical energy, but must have been derived from some other source. In other words, the electrical energy of the cell when it was sending a current must have been greater than the chemical energy corresponding to the chemical changes which took place during the passage of this current.

Next suppose that the E.M.F. of the cell decreases as the temperature increases, so that when the temperature is T_1 the E.M.F. is represented by OD, while when the temperature is T_2 the E.M.F. is OA. Then, if when the cell is at a temperature T_1, it is allowed to send a current till Q units of electricity have passed, the work done will be represented by DCMO, while the work which must be done to drive Q units in the reverse direction when the temperature is lowered to T_2 is represented by ABMO. Hence in this case more work has to be done by the external source than is done by the cell when such a cycle is traversed, that is, the electrical energy which the cell supplies is not as great as would be expected from thermo-chemical data.

Since energy can neither be created nor destroyed, it follows that in the first case considered, namely, when the E.M.F. of the cell decreased with decrease of temperature, since more work is done by the cell when sending the current than is supplied by the chemical changes which take place, the extra energy will have to be supplied at the expense of the heat of the cell, so that if no outside heat is supplied the cell will get cooler as the current passes, or to keep it at a constant temperature heat must be supplied. In the second case, where increase of temperature causes decrease of E.M.F., the opposite is the case, and the cell will get hotter when sending a current, and to keep its temperature constant heat must be abstracted.

By means of the second law of thermo-dynamics it is possible to calculate the quantity of heat which must be supplied or abstracted in this way. Consider the first case, where the temperature coefficient of the cell is positive, that is, where increase of temperature is accompanied by increase in the E.M.F. Here heat has to be supplied while the cell is passing from the condition represented by the point A to that represented by B, and abstracted, since we are now sending the current in

the reverse direction, while passing from C to D. Let H_1 be the heat (measured in ergs) supplied at the temperature T_1, and H_2 the heat abstracted at the temperature T_2, then, by the second law of thermodynamics (§ 261),

$$\frac{H_1 - H_2}{H_1} = \frac{T_1 - T_2}{T_1}.$$

But $H_1 - H_2$ is the heat used during the cycle, and is equal to the rectangle ABCD, which we have seen is equal to $Q(E_1 - E_2)$. Hence

$$\frac{Q(E_1 - E_2)}{H_1} = \frac{T_1 - T_2}{T_1},$$

or

$$H_1 = QT_1 \frac{E_1 - E_2}{T_1 - T_2}.$$

But $\frac{E_1 - E_2}{T_1 - T_2}$ is the rate of change of the E.M.F. of the cell with temperature, that is, the temperature coefficient, so that if we represent this by $\frac{\delta E}{\delta T}$, and if, further, we take Q as one unit, we have that the quantity of heat h converted into electrical energy during the passage of the unit quantity of electricity is

$$h = T\frac{\delta E}{\delta T}.$$

If the E.M.F. decreases with increase of temperature, so that the temperature coefficient is negative, we have simply to change the sign of $\frac{\delta E}{\delta T}$.

Hence if h is the total quantity of heat produced by the chemical changes which go on in a cell during the passage of unit quantity of electricity, and E is the E.M.F. of the cell, the following equation will hold :—

$$h + T\frac{\delta E}{\delta T} = E.$$

Thus, in order to be able to calculate the E.M.F. of a cell from thermochemical data, it is necessary to know the temperature coefficient of the cell.

Since in the case of the Daniell cell the E.M.F. calculated without taking account of the effect of the temperature coefficient agrees with the observed value, it is evident that in the case of this cell the E.M.F. can only vary very little with temperature, and experiment has shown that this is the case, the temperature coefficient being +0.000034.

559*. Experimental Verification of the Helmholtz Formula.—The direct experimental verification, by means of thermal measurements, of the correctness of Helmholtz's expression for the difference

between the electrical energy available in a cell and the chemical energy corresponding to the processes which go on in the cell during the time it is sending a current, has been undertaken by Jahn. The cell to be examined was placed within a Bunsen's ice-calorimeter (§ 212), while there were two external circuits each containing a galvanometer. One of these circuits, AG_1B (Fig. 531), was of low resistance, and the other, AG_2B, of very high resistance. This being so, practically the whole of the current sent by the cell passes through the galvanometer G_1, and this reading of the galvanometer serves to measure the current passing. The very small current which passes through the high resistance galvanometer, G_2, is proportional to the difference in potential between the points A and B, so that the deflection of this galvanometer serves to measure the difference of potential between A and B.

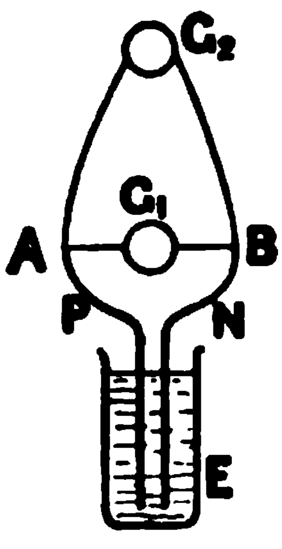


FIG. 531.

Let C be the current sent by the cell and e the difference in potential between the points A and B when this current is passing. Of course e will be less than the E.M.F. E of the cell on open circuit, since the cell itself has resistance as well as the wires connecting the poles to the points A and B, and therefore, according to Ohm's law, there will be a fall of potential when a current is passing. The heat developed in the branch AG_1B in the time t will be equal to eCt in electrical units or $aeCt$ in calories, where a is the value of one joule in calories, that is, 0.2387. If r_1 is the resistance of the wires connecting the terminals of the cell with the points A and B, between which e is measured, the energy spent in these wires during a time t will be by Joule's law r_1C^2t joules, or ar_1C^2t calories. Hence the total energy expended by the cell on the portions of the circuit outside the calorimeter is $aC(e+Cr_1)t$.

In addition to the heat developed in the external circuit of the cell, there will be heat developed within the cell itself, owing to the passage of the current through the electrolyte, and if r_2 is the resistance of the cell, the quantity of heat developed in this way will be ar_2C^2t calories. The total heat developed in the circuit will thus be

$$aC(e+Cr_1)t+ar_2C^2t,$$

and this represents the total quantity of energy transformed during the passage of Ct units of electricity through the cell. We have seen in § 550 that if E is the E.M.F. on open circuit of a cell of which the internal resistance is r_2, then the E.M.F. between the terminals, when it is sending a current C, is $E-r_2C$, so that the difference of potential between the points PN (Fig. 531) is $E-r_2C$, where E is the E.M.F. of the cell measured on open circuit. Since the resistance of the two wires PA and BN is r_1, there will be a further decrease in the difference

of potential, equal to r_1C, so that the difference of potential between the points A and B will be given by

$$e = E - C(r_1 + r_2).$$

Hence $$E = e + C(r_1 + r_2),$$

and therefore $$aC(e + Cr_1)t + ar_2C^2t = aCET \quad . \quad . \quad . \quad (1).$$

Each side of this equation represents the total quantity of energy directly or ultimately transformed into heat when Ct units of electricity pass through the cell. The total quantity of energy transformed into heat outside the calorimeter is $aC(e + Cr_1)t$. Now we can observe the quantity of heat actually evolved in the calorimeter, but this heat will not be equal to ar_2C^2t, because, according to Helmholtz's theory, there is a quantity of heat $aCT\frac{\delta E}{\delta T}.t$ absorbed from the environment when a quantity of electricity Ct passes at constant temperature T through a reversible cell of E.M.F. E, if $\frac{\delta E}{\delta T}$ represents the rate at which the E.M.F. of the cell increases with increase of temperature. Thus W, the actually observed quantity of heat evolved in the calorimeter, is given by

$$W = ar_2C^2t - aCT\frac{\delta E}{\delta T}.t,$$

and therefore by substitution in equation (1) we get

$$aC(e + Cr_1)t + W + aCT\frac{\delta E}{\delta T}t = aCEt.$$

Since $aC(e + Cr_1)t$ represents the quantity of energy, expressed in thermal units, converted into heat in the portion of the circuit outside the calorimeter, if Q is taken to represent the total quantity of energy, in thermal units, which passes from the cell and circuit to its surroundings during the passage of Ct units of electricity, we have

$$Q + aCT\frac{\delta E}{\delta T}t = aCET,$$

or $$aCT\frac{\delta E}{\delta T}t = aCET - Q \quad . \quad . \quad . \quad (2).$$

Hence, by measuring the current strength C, the difference of potential between A and B and the heat, W, developed in the calorimeter, Q can be calculated, and then, knowing the E.M.F. E of the cell on open circuit, by means of equation (2), the value of the term $aT\frac{\delta E}{\delta T}Ct$ can be calculated. Since the cell is placed in an ice-calorimeter, its temperature will be 0° C. or 273° on the absolute scale, so that T is 273. The quantities a, C, and t being known, the value of the temperature coefficient $\frac{\delta E}{\delta T}$ can be calculated, and the value thus obtained from

thermal measurements can be compared with the value found by direct electrical measurements, and the agreement or otherwise between these values will be evidence as to the accuracy of Helmholtz's theory.

In an experiment, using a Daniell cell in which the concentration of the electrolytes was in each case 1 gram molecule of the salt to 100 gram molecules of water, Jahn found that E, the E.M.F. on open circuit, was 1.096 volts at 0° C. Hence, if the quantity of electricity Ct which passes is one coulomb, the term

$$\begin{aligned} aECt &= .2387 \times 1.096 \\ &= .2617 \text{ calories.} \end{aligned}$$

In the calorimetric experiment the current was allowed to pass for an hour, and the mean value of the current was 0.066573 ampere. The mean value of the product of the current into the difference of potential, e, between the points A and B was 0.011222. During the course of the experiment the motion of the mercury thread of the calorimeter indicated that 52.394 calories had been developed, so that is the value of W. Hence, noting that the resistance r_1 of the wires AP and BN was 0.1 ohm, the following quantities of heat were developed :—

$$\begin{aligned} aCet &= 9.659 \text{ calories.} \\ ar_1C^2t &= 0.381 \quad \text{,,} \\ W &= 52.394 \quad \text{,,} \end{aligned}$$

Hence $$Q = 62.434.$$

Since the mean value of the current was 0.066573 ampere, and it flowed for 1 hour or 3600 seconds, the total quantity of electricity which flowed through the cell was 0.066573 × 3600 or 239.66 coulombs.

If one coulomb had passed through the cell the heat developed in the whole of the circuit would have been

$$\frac{62.434}{239.66} = .2604 \text{ calories.}$$

Hence, when the quantity of electricity, Ct, which passes through the circuit is one coulomb, we have

$$\begin{aligned} aT\frac{\delta E}{\delta T}Ct = aCEt - Q &= 0.2617 - 0.2604 \\ &= 0.0013 \text{ calories.} \end{aligned}$$

Therefore, since T is 273 and a is 0.2387, we get

$$\frac{\delta E}{\delta T} = \frac{.0013}{.2387 \times 273} = 0.00002 \text{ volts per degree.}$$

The value of the temperature coefficient of the Daniell got from direct electrical measurements is 0.000047, a number which agrees very well with that obtained above from the calorimetric measurements, on the supposition that Helmholtz's theory is correct. Although 20 and

47 may seem very different, it must be remembered that the coefficient obtained is practically zero, and such differences as appear are quite within the limits of experimental error.

As another example, we may take the cell consisting of a plate of silver surrounded with silver chloride and a plate of zinc in a solution of zinc chloride. The zinc chloride solution contained one gram molecule of $ZnCl_2$ in 50 gram molecules of water.

The E.M.F. E on open circuit of this cell is 1.0171 volts. Hence, when one coulomb passes through the cell, we have

$$aECt = .2387 \times 1.0171 = 0.2428 \text{ calories.}$$

The experiment lasted an hour, the mean current being 0.093041 amperes, and the mean value of the product eC being 0.0233669 joules. The heat liberated in the calorimeter during the experiment was 64.192 calories. Hence

$$\begin{aligned} aCet &= 20.339 \\ ar_1C^2t &= .744 \\ W &= 64.192 \\ \hline \therefore \quad Q &= 85.275 \end{aligned}$$

The total quantity of electricity which passed through the cell was $0.093041 \times 3600 = 334.94$ coulombs. Hence if one coulomb had passed, the total quantity of heat developed in the circuit would have been

$$\frac{85.275}{334.94} = 0.2546 \text{ calories.}$$

Hence $$aT\frac{\delta E}{\delta T}Ct = 0.2428 - 0.2546 = -0.0118 \text{ calories.}$$

$$\therefore \quad \frac{\delta E}{\delta T} = -\frac{0.0118}{.2387 \times 273} = -.00018 \text{ volt per degree.}$$

The temperature coefficient of this cell obtained by direct measurement is 0.00021 volt per degree, so that the agreement is very satisfactory.

When a lead accumulator sends a current the sulphuric acid solution is electrolysed, the hydrogen ions go to the positive plate (the lead peroxide plate), and the SO_4 ions to the negative plate, that is the metallic lead plate. The hydrogen ions act on the lead peroxide according to the equation

$$PbO_2 + H_2 = PbO + H_2O + 58300 \text{ calories.}$$

That is during this reaction, in which one gram molecule of the compounds take part, there is an evolution of heat of 58300 calories. Next the sulphuric acid acts on the lead oxide to form lead sulphate according to the equation

$$PbO + H_2SO_4 = PbSO_4 + H_2O + 23400 \text{ calories.}$$

Thus the total thermal value of the changes of the positive plate is 81700 calories.

At the negative plate the SO_4 ions react with the water to form H_2SO_4 and O. Now the passage of the current has split sulphuric acid into H_2 and SO_4, and then this SO_4 has by combination with water reformed sulphuric acid so that the whole reaction amounts to the splitting up of water into hydrogen and oxygen, for just as much heat will be given out when the acid is reformed as was absorbed when it was split up. Hence we have to include in our thermal equations the splitting up of water into hydrogen and oxygen, that is,

$$H_2O = H_2 + O + 68400 \text{ calories.}$$

The oxygen combines with the lead to form lead oxide, and

$$Pb + O = PbO + 50300 \text{ calories.}$$

This lead oxide is then converted into lead sulphate, and

$$PbO + H_2SO_4 = PbSO_4 + H_2O + 23400 \text{ calories.}$$

Thus the total thermal value of these changes is 5300 calories. Hence for the whole cell we have that the thermal value of the chemical changes which take place is

$$81700 + 5300 \text{ or } 87000 \text{ calories.}$$

This is the quantity of heat which corresponds to the electrolysis of one gram *molecule* of sulphuric acid, or the liberation of two gram equivalents of hydrogen. This corresponds to the passage of 96550×2 coulombs. Also since one joule is equal to 0.2387 calories the mechanical value of the thermo-chemical changes which go on in the cell is 87000/0.2387 joules.

If V is the E.M.F. of the cell, then the passage of 96550×2 coulombs corresponds to the performance of $96550 \times 2V$ joules of work. Thus, assuming that the temperature coefficient of the cell is zero, we have the equation

$$96547 \times 2V = 87000/0.2387,$$

$$V = \frac{87000}{96547 \times 2 \times 0.2387}$$

$$= 1.888$$

Hence the E.M.F. corresponding to the chemical changes is 1.888 volts. Experiments have shown that the temperature coefficient of a cell made with solution of the concentration of that used in the thermal measurements is +0.00014. Hence at a temperature of 15° C., or 288° absolute,

$$T\frac{\delta E}{\delta T} = 288 \times 0.00014 = .040.$$

Thus the E.M.F. of such a cell at 15° C. is 1.888 + 0.040 volts, or 1.928 volts. An experimental measurement of the E.M.F. gave 1.900, a number which is in very fair agreement with that deduced from the

thermal data. The agreement between the number calculated and the observed value is also a strong argument in favour of the accuracy of the account which has been given above of the chemical changes which take place in the accumulator.

560. Heat developed in a Circuit when the Current performs Mechanical Work.—We have hitherto considered that the whole of the energy of the current has been spent either in the production of heat in the circuit, or in the performance of chemical work. The current may, however, do mechanical work; for instance, it may drive an electric motor, and this motor may be employed in raising a weight against gravity, or pumping water from a lower to a higher level. In such a case, some of the energy being converted into potential energy in the raised weight or water, the amount of heat developed in the circuit during a given amount of chemical change in the cell, will not be so great as when no external work is done.

Let us consider the case of a cell, say an accumulator, for in this case the internal resistance of the cell is so small that it may be neglected, connected to an electric motor. Let the resistance of the circuit, including the motor and the leads, be R, and the E.M.F. of the cell, E. In the first place, let the armature of the motor be fixed so that it cannot rotate, and therefore none of the electrical energy will be converted into energy of motion of the motor. If C is the current which flows through the circuit under these circumstances, then by Ohm's law $C=E/R$, and the heat developed in the circuit in a time t is equal to tC^2R. Now Ct is the quantity of electricity Q which passes round the circuit, so that the heat developed is RCQ or EQ. Next suppose that the armature is released, but that it is allowed to turn freely, so that all the energy supplied to the motor is employed in overcoming the friction of the different parts of the motor, and in the production of heat in the armature and field magnet coils. Let e be the back E.M.F. produced in the armature (§ 530) when the motor has reached a steady state, so that it is rotating at a uniform speed. Then the effective E.M.F. in the circuit is $E-e$. Hence the current which passes is now $(E-e)/R$. The heat developed in the wire constituting the leads, the armature, and the field-magnet coils during the passage of Q coulombs will now be $Q(E-e)$. In addition, since a quantity of electricity Q is forced against an E.M.F., e, an amount of work will have to be done represented by Qe. Since the motor is doing no work, the energy represented by Qe will simply be frittered away as heat due to friction of the different moving parts, so that the total amount of heat produced in the circuit during the passage of Q coulombs will be $Q(E-e)+Qe$ or QE. Hence, as before, the whole of the electrical energy derived from the cell is converted into heat.

Next suppose that the motor is employed in the performance of external work, say the raising of a weight, and that during the passage of Q coulombs of electricity the work done is W. If e is the back E.M.F.

in this case, the effective E.M.F. for sending a current in the circuit is $E-e$, and hence the heat developed in the circuit, according to Joule's law, is $(E-e)Q$. In addition to the work done in raising the weight, which will be stored up as potential energy, a certain amount of work will have to be done to overcome the friction at the bearings, &c., of the motor and the mechanism which is used to raise the weight; this work will appear as heat developed owing to the friction, and if h is the amount of such work which corresponds to the passage of Q coulombs of electricity through the circuit, we have, since the total work done by the cell must be equal to the product of the E.M.F., E into the quantity of electricity which is sent through the circuit by the cell, that is, QE,

$$QE = Q(E-e) + W + h.$$

Here $(E-e)Q+h$ represents the quantity of energy which is converted into heat, and W is the energy which is stored up as potential energy due to the weight which has been raised.

561*. Heat of Ionisation.—We have seen that the relation between the E.M.F. of a reversible cell, the chemical energy of separation transformed, and the heat taken from the surroundings of the cell while working at constant temperature, is expressed by an equation of the form

$$E = h + T\frac{\delta E}{\delta T}$$

in which h is the loss in intrinsic "chemical" energy of the system when unit quantity of electricity passes, and $T\frac{\delta E}{\delta T}$ is the corresponding quantity of heat absorbed from the environment, according to Helmholtz's theory. Recently a very interesting attempt has been made to analyse still further the processes which go on in the cell. We shall briefly indicate the nature of this analysis, taking for the purpose of illustration the case of the reversible Daniell cell.

When this cell is allowed to produce 2×96550 coulombs of electricity by closing the circuit, 65 grams of Zn are dissolved from the Zn electrode and 63 grams of Cu are deposited on the Cu electrode. The change in intrinsic energy which here occurs is obviously the same as that which takes place, when, by the introduction of a zinc plate into a solution of copper sulphate, 63 grams of copper are precipitated. The heat evolved during this latter process has been observed to be about 501,000 calories.

If we represent by $2e_0$ the quantity of electricity passing when 65 grams of Zn are dissolved. and 63 grams of copper precipitated, then

$$2e_0E = H + 2e_0T\frac{\delta E}{\delta T},$$

where H represents an amount of energy equal to 501,000 calories. This quantity, H, is usually taken to represent the loss of "chemical" energy by the system, but no attempt is made to analyse it further. On the ionic

hypothesis, the following is what has occurred: 65 grams of Zn have passed from the neutral state into the solution, where they exist as ions, and 63 grams of Cu, originally existing in the solution as ions, have passed into the state of ordinary, electrically neutral copper. We may represent the change as

$$\mathrm{Zn} + \mathrm{Cu}^{\cdot\cdot}\,\mathrm{SO_4}''\mathrm{Aq} = \mathrm{Zn}^{\cdot\cdot}\,\mathrm{SO_4}''\mathrm{Aq} + \mathrm{Cu} + H,$$

where Zn and Cu signify electrically neutral equivalent weights of the respective metals, possessing the amounts of energy corresponding to these states, and $\mathrm{Cu}^{\cdot\cdot}$ and $\mathrm{Zn}^{\cdot\cdot}$ signify the same weights of Cu and Zn, with the amounts of energy which they possess when ionised. The symbol $\mathrm{SO_4}''$ refers to the ionised acid radicle, and the equation assumes that the degree of dissociation of the $ZnSO_4$ solution is the same as that of the $CuSO_4$ solution.

It is clear that we may regard the change as consisting in the ionisation of an equivalent weight of Zn, and the passage from the ionic to the neutral state of an equivalent weight of Cu. Now, we may suppose that there is some definite relation between the energy of a given mass of an element in the neutral state and the energy of the same mass when in the ionic state. If the energy in the ionic state is the smaller, then ionisation will be accompanied by development of heat, which we may call the "heat of ionisation." It is evident that H in the above equation signifies, following this idea, the difference between the respective heats of ionisation of Zn and Cu. Thus we may write

$$\mathrm{Zn} = \mathrm{Zn}^{\cdot\cdot}.\mathrm{Aq} + h_1$$

and

$$\mathrm{Cu} = \mathrm{Cu}^{\cdot\cdot}.\mathrm{Aq} + h_2,$$

where h_1 and h_2 represent the heats of ionisation of Zn and Cu respectively, and are connected by the relation

$$H = h_1 - h_2.$$

The question now arises whether there is any means by which we may obtain the separate heats of ionisation, and not their difference merely. The expression

$$2e_0E = H + 2e_0T\frac{\delta E}{\delta T}$$

is a particular case of a general theorem which can be applied to the changes that occur at each electrode-electrolyte surface. Thus, if E_1 represents the amount by which the potential of the solution exceeds that of the Zn electrode, and E_2 that by which the copper exceeds that of the solution in its neighbourhood, we have

$$2e_0E_1 = h_1 + 2e_0T\frac{\delta E_1}{\delta T},$$

and

$$2e_0E_2 = -h_2 + 2e_0T\frac{\delta E_2}{\delta T}.$$

On addition these equations give

$$2e_0(E_1+E_2)=(h_1-h_2)+2e_0T\frac{\delta(E_1+E_2)}{\delta T},$$

and are thus seen to be equivalent to the original Helmholtz equa-
(§ 558). The quantities $2e_0T\frac{\delta E_1}{\delta T}$ and $2e_0T\frac{\delta E_2}{\delta T}$ represent the respec-
quantities of heat absorbed at the electrode-electrolyte surface w-
65 grams of Zn pass into solution and when 63 grams of Cu pass o-
of solution. The values of these quantities can be estimated from obs-
vations upon the thermal effects at the electrodes, when a salt of ea-
metal is electrolysed between electrodes of the same metal. In order
determine h_1 and h_2 it is further necessary to know the values of E_1 a-
E_2. If we suppose these to be known from capillary electrometer da-
we can finally determine h_1 and h_2.

Thus for copper we get

$$h_2 = -17700 \text{ calories},$$

and for zinc $$h_1=33100 \text{ calories}.$$

And hence $$h_1-h_2=50800 \text{ calories},$$

which result agrees sufficiently nearly with the observed value,

$$H=50100 \text{ calories}.$$

The results may also be expressed

$$Zn=Zn^{\cdot\cdot}+33100$$
$$Cu=Cu^{\cdot\cdot}-17700.$$

It would thus appear that the intrinsic energy of ionised Zn is les-
than that of an equal mass of electrically neutral Zn, while the opposi-
is true in the case of copper. The tendency of the Zn to ionise and o-
the Cu to become neutral, as exhibited during the working of the Danie-
cell, follows naturally from these results.

Considerations of this kind appear to mark a distinct advance toward-
a more complete physical theory of chemical change.

CHAPTER XVIII

PASSAGE OF ELECTRICITY THROUGH GASES

562. Passage of Electricity through Gases.—We have considered the chief phenomena which accompany the passage of electricity through metallic conductors and electrolytes, and we have now to consider the passage of electricity through gases.

In the case of gases it will be necessary to consider the pressure to which the gas is subjected when the passage of the electricity takes place, for the phenomena vary enormously as the pressure is altered.

Suppose we have two conductors which are separated by a gas, say air, at atmospheric pressure, and the difference of potential between the conductors is gradually increased. Then owing to this difference of potential, the air between the conductors will be put into a condition of electrical strain.

As the potential difference is increased the strain on the air will increase till a condition will be reached when the air is no longer able to support the strain, and it breaks down and allows a current to pass. This forms what is called the spark discharge. Before the passage of the spark there will be a fall of potential in the air between the conductors, and the fall of potential per unit length is a measure of the electrical stress tending to break down the gas, or as it is called, the electromotive intensity acting on the gas. The maximum electromotive intensity which a gas can support before a spark passes has been called by Maxwell the electric strength of the gas. Experiment has, however, shown that the electric strength thus defined depends on a number of conditions besides the nature of the gas and the pressure to which it is subjected. Thus the electromotive intensity is found to depend to a small degree on the nature of the conductors between which the spark is passed, and to a greater degree on the shape of the surfaces of these conductors between which the spark passes. The electromotive intensity also depends very much on the distance between the conductors.

This effect is very clearly shown by the numbers given in the following table, which shows the potential difference (measured in electrostatic units) required to cause a spark between two spheres of 9.76

cm. radius when they are separated by different distances in air at a pressure of 76 cm. of mercury:—

Spark Length in cm.	Potential Difference (Electrostatic Units).	Electromotive Intensity.
.0066	2.63	399
.011	3.36	320
.1	15.00	150
.56	63.70	114
1.07	110.78	104

It will be observed that the electromotive intensity necessary to produce a spark, when the distance between the spheres is small, is very much greater than when this distance is comparatively great. Such results as these show that the electrical strength of a gas is not a property of the gas alone, but is a complex quantity depending on a number of considerations besides the properties of the gas.

If the pressure of the gas is altered, the difference of potential necessary to produce a spark varies very greatly. As the pressure is reduced from the atmospheric pressure the difference of potential required to produce a spark of a given length decreases at first; but this decrease

Fig. 532.

does not go on indefinitely, for a critical pressure will eventually be reached, and when the pressure is further decreased the potential necessary to produce a spark will increase, and this increase will go on as long as the pressure is decreased, so that at the highest vacua attainable the gas will be a perfect insulator, and it will be impossible to pass a spark. The critical pressure, at which the electromotive intensity is a minimum, varies with the distance between the electrodes. Thus while for a spark length of 1/100 mm. the critical pressure in air is equal to the pressure of 25 cm. of mercury, for a spark length of several millimetres in length the critical pressure is less than that due to a millimetre of mercury.

The appearance of the discharge changes in a very marked manner as the pressure of the gas is reduced. Thus at atmospheric pressure the spark consists of a brilliant line of light which is sharply defined, and is either straight if the spark length is small or is bent in a very characteristic manner, being sometimes of a forked nature.

Suppose that a glass tube, such as is shown in Fig. 532, with platinum wires fused through the ends, these wires being connected with a small aluminium plate K and a wire A, is gradually exhausted, and that a discharge is passed through the gas in the tube, the plate A being the positive, or, as we may call it, the anode, and K the kathode. When the pressure is equal to about 8 cm. of mercury there will be a line of light stretching down the axis of the tube somewhat as shown in the figure. If the pressure is reduced to about half a millimetre of mercury, then the general appearance when the discharge passes is that shown in Fig. 533. At the kathode K there will be seen a soft glow which moves about over the surface of the electrode. Next to the kathode there is a space, B,

Fig. 533.

which is comparatively free from luminosity, and which is called Crookes's space, or the first dark space. The distance from the kathode through which this dark space stretches increases as the exhaustion of the gas increases. The termination of the dark space nearest the anode is quite sharp, and is very approximately the surface on which would lie the ends of equal normals drawn from the surface of the kathode. Beyond the dark space is a luminous space C, called the negative column. The position of the negative column does not depend on that of the anode, so that if the anode is placed in a side tube the negative column does not bend round into the side tube, but goes straight on and fills the portion of the tube beyond the point where the side tube containing the anode leaves the main tube.

Beyond the negative column there is a second comparatively dark space D, called the second negative dark space. This dark space varies very much in size, and may sometimes be entirely absent. Beyond this dark space there is another luminous column, E, which extends up to the anode, and is called the positive column. The luminosity of the positive column is often not continuous, but consists of alternate bands of bright

light and comparatively dark spaces. These bright bands are called striæ, and often present a very striking appearance. The colour of the striæ depends on the nature of the gas within the tube, and, according to Crookes, when a mixture of gases exists within the tube each gas produces a separate series of striæ.

While the negative column and the dark space are confined to the neighbourhood of the kathode, and do not increase in size if the distance between the anode and kathode is increased, the positive column always stretches up to the anode, passing along the shortest path from the kathode to the anode. Professor J. J. Thomson has passed a discharge through an exhausted tube 50 feet in length, and with the exception of a few inches near the kathode the positive column filled the whole of the tube and exhibited very well marked striations throughout its length. It is probable that the discharge is actually carried by the positive column, and that the other phenomena observed near the kathode are simply due to some peculiarities which seem always to accompany the passage of electricity from a gas to a conductor.

When the exhaustion of the tube is carried considerably below that for which the discharge has the appearance just described, the character of the discharge is quite altered, and a series of subsidiary phenomena occur which, especially of late years, have attracted much attention. It will be convenient to consider these phenomena at very high vacua in a separate section.

563. Kathode Rays.—When the exhaustion within a tube, such as is shown in Fig. 534, is carried to below a thousandth of a millimetre of mercury, the positive column gradually vanishes, and the sides of the

FIG. 534.

tube exhibit a brilliant phosphorescent glow. The colour of this glow depends on the nature of the glass, thus with lead, or English glass, the glow is blue, while with German, or soda glass, the phosphorescence is of a beautiful emerald green. The appearance presented is as if something were projected by the kathode in a direction normal to its

surface which, when it strikes the glass, has the power of exciting phosphorescence.

The phenomena in these high vacua have been studied at great length by Crookes, who supposes that the comparatively few molecules of the gas which are left in the tube become electrified by contact with the kathode, and that these negatively electrified gas molecules are then shot out by electro-static repulsion from the kathode. He further supposes that at these high degrees of exhaustion the number of molecules is so small that the molecules will travel for considerable distances without encounters one with the other, so that the molecules which are shot out from the kathode do not lose their energy of translation by sharing it with other molecules, with which they would collide if the density of the gas were greater, but reach the walls of the tube while still moving with a high velocity, and there by their impact develop phosphorescence in the glass.

Whatever the nature of the emanation from the kathode, or the kathode rays as they are called, they under ordinary circumstances proceed in straight lines, so that if a screen, such as is shown in Fig. 535,

FIG. 535.

is placed between the kathode and the sides of the tube a clear shadow of the screen will be produced on the walls of the tube, no phosphorescence taking place within the portion of the wall in the geometrical shadow. It is immaterial in this experiment whether the screen be composed of a conductor or a dielectric. If, instead of a plane plate, a concave plate is used for the kathode, the kathode rays are brought to a focus at the centre of curvature of the kathode, and if a body is placed at this focus it will be raised to a bright incandescence, or, if the discharge is sufficiently strong, even melted owing to the impact of the rays. In this way platinum can be melted. Other substances besides glass phosphoresce when the kathode rays are allowed to fall on them. Thus if some rubies are enclosed in a vacuum tube and the discharge

is passed, the kathode rays falling on the rubies will cause them to give out a brilliant ruby-red light.

The kathode rays are deflected when a magnet is brought near in the direction which would occur if they formed a flexible metallic conductor conveying the current from the anode to the kathode. Thus if a screen pierced with a slit is placed before the kathode as shown in Fig. 536, so that a narrow beam of the kathode rays is obtained, when no magnet is near these rays will travel straight on and strike the glass at C, where they will cause phosphorescence. When a magnet is brought near the tube so that the lines of force are at right angles to the paper, the rays are deflected and the phosphorescent patch on the glass is moved down into the position D. Not only will the phosphorescent patch be deflected, but while when undeflected it was a continuous patch, when deflected it will consist of a number of bright bands separated by more or less dark spaces. Thus under the influence of the deflecting magnetic field not only do the rays get deflected, but it would appear that a species of dispersion is also produced, so that all the rays do not get deflected to the same extent, and, just as the spectrum of an incandescent gas consists of a certain number of bright lines, so does this kathode ray "spectrum," produced by the action of a magnetic field, also consist of bright bands.

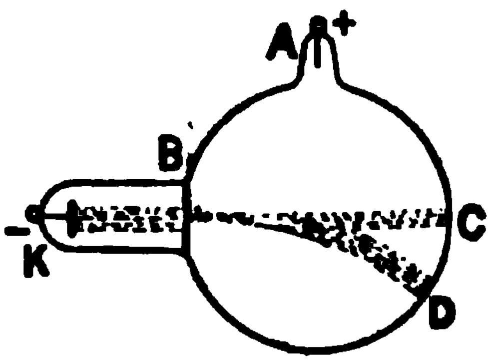


FIG. 536.

Crookes' theory that the kathode rays consist of negatively charged particles shot out from the neighbourhood of the kathode is not accepted by some observers who suppose that the kathode rays are really ether waves of the same general nature as light waves but of different wavelength, so that their effects cannot be observed by the ordinary means used to study light waves. The fact that the kathode rays are deflected by a magnet is, however, a strong argument against this view, since no one has been ever able to detect that a magnetic field has any deflecting effect on light waves. In addition it can be shown experimentally that along the path of the kathode rays there is a transport of negative electricity. The experiment was originally devised by Perrin, and has been modified by J. J. Thomson. A tube is taken of the form shown in section in Fig. 537, in which K is the kathode and A is the anode. In a side branch a metal tube, C, with a slit at the end is placed and connected to earth. Inside this tube, but insulated from it, is another metal tube, D, which also has a slit which is opposite the slit in the outer tube. This tube is connected with an electrometer. When the discharge is passed the kathode rays travel straight across the tube and strike the glass at the point B and the electrometer is undeflected. If, however, by means of a magnet the kathode rays are made to enter the slit in the tubes C and D the electrometer will be deflected, showing that

the tube D has acquired a negative charge. The outer tube C which is connected with the earth will screen off from the inner tube all electrical disturbances other than such as enter by the slit. When the negatively charged particles enter the inner tube, since they are at the inside of a conductor, they will either give up their charges to the conductor or will induce a positive charge on the inside of the conductor, and the corresponding negative charge will cause the electrometer to be deflected. Hence the deflection of the electrometer seems clearly to indicate that the kathode rays are at any rate always accompanied by the projection of negatively charged particles, and hence it is only natural to suppose that they really consist of such negatively charged particles.

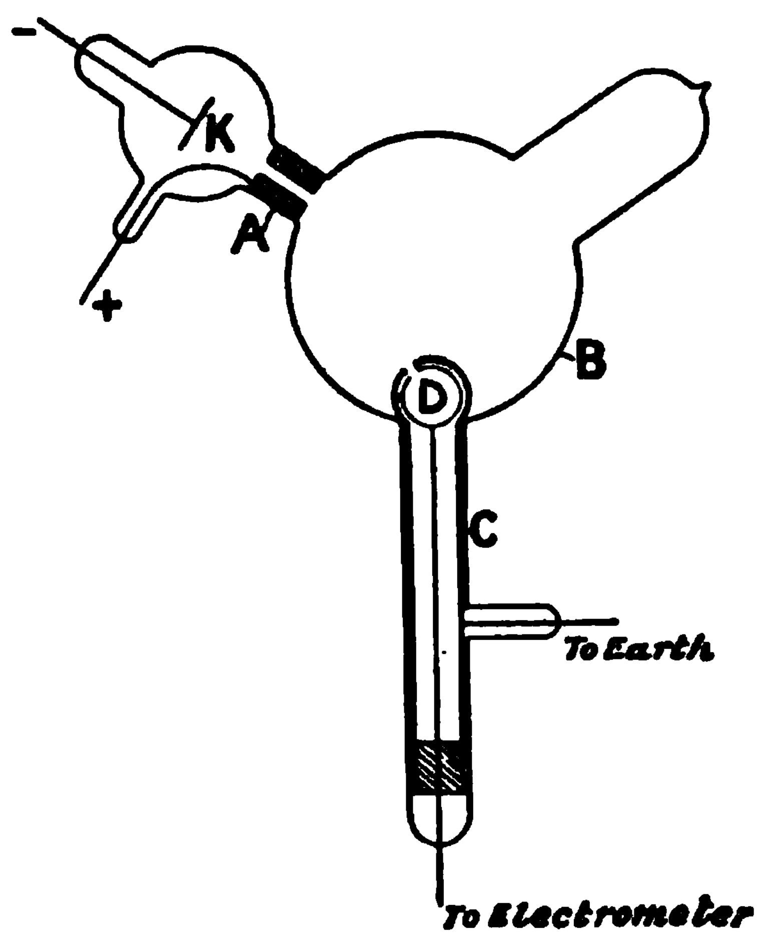


FIG. 537.

We have hitherto only considered the behaviour of kathode rays within the highly exhausted tube where they are produced. Lenard has, however, found that the kathode rays can be obtained in the space outside the tube. Thus if a small window in the tube is covered with very thin aluminium, and the kathode rays are directed on to this window, rays are found to proceed from this window which are deflected by a magnet and are capable of producing phosphorescence in the same way as do the kathode rays within the tube. These rays are also found to carry a negative charge. At first this possibility of passing the rays through a sheet of aluminium seems a very strong argument against the theory that they consist of negatively electrified particles. It has, however, been found that if a piece of metal inside a Crookes' tube is placed in the path of the kathode rays it acquires the property of itself giving off kathode rays, as if the shock due to the impact of the kathode rays were capable of driving off the gas[1] particles which are in contact with the metal. Thus the kathode rays outside the tube in Lenard's experiment are not due to the motion of the same particles which struck the inside of the aluminium plate, but consist of a new set of electrified particles which are shot out from the outside surface

[1] It must be noted that, according to J. J. Thomson, the mass of the particles which carry the negative charges are considerably less than the mass of the molecules of the gas, so that we seem here to be dealing with something smaller than the molecule.

of the plate, owing to the impact of the charged particles which strike the inside surface.

564. Röntgen Rays.—When the kathode rays strike upon matter, in addition to inducing phophorescence and raising the temperature and in certain cases causing the emission of kathode rays, there is produced a kind of rays which differ from the kathode rays and from ordinary light rays in very many particulars. These rays were first discovered by Röntgen, and are therefore called Röntgen rays or X rays. They differ from kathode rays in that they pass through glass and many other materials with comparatively little absorption, and they are not deflected by a magnet as are the kathode rays. They differ from ordinary light rays in that they do not appear to be refracted when they pass from one medium to another, while it has been found impossible to obtain by their means any signs of interference phenomena.

The form of tube which has been found best for producing these rays is shown in Fig. 538. It consists of a concave aluminium kathode K and a platinum plate, A, as anode, which is inclined at 45° to the axis of the kathode. The tube itself is formed of soda glass, since it is found that lead glass is very opaque to the rays. The kathode rays are focussed on the platinum anode and the Röntgen rays proceed as if they came in straight lines from the point of the anode where the kathode rays strike it.

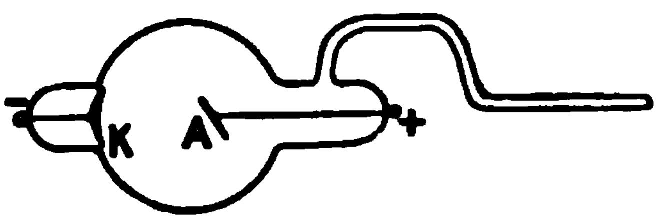


FIG. 538.

The presence of Röntgen rays can be detected either by their action on a photographic plate or by the fluorescence which they excite when they fall on some substances, such as the double cyanide of potassium and platinum. The most striking peculiarity of these rays is that they are capable of penetrating many substances which are opaque to ordinary light. Thus black paper, wood, and aluminium are transparent, while the more dense metals, such as lead, are opaque to the rays. The most important practical application of the differences between the transparency of different bodies to these rays is owing to the fact that while flesh is fairly transparent the bones are very much more opaque. Thus if a tube producing Röntgen rays is placed above the hand while a photographic plate is placed below the hand, the rays will pass through the flesh of the hand and will act on the plate. The bones, however, will stop the rays, and hence those parts of the plate within the shadow of the bones will only be slightly affected, and on developing the plate a shadow of the bones of the hand will be obtained. Such a photograph is shown in Fig. 539, and it will be observed that the bones are much more opaque than the flesh. Instead of using a photographic plate, a paper screen which is coated with one of the salts which fluoresces when

the Röntgen rays fall on it may be used, when the salt will fluoresce where the rays are transmitted by the flesh, but not where the rays have been absorbed by the bones, and so a dark shadow of the bones will appear on the screen.

No conclusive experimental evidence as to the nature of the Röntgen rays has yet been obtained. It is, however, probable that they are of the nature of a wave disturbance in the ether, the wave-length being very small compared with that of blue light. J. J. Thomson has shown that waves which would have the same properties as the Röntgen rays, in that they would not be refracted or show interference, would be produced in the ether if we suppose that the kathode rays are really negatively charged particles which strike the body which is producing the rays. The disturbance produced would not be so much a wave motion as a single wave or impulse which would travel out through the medium.

Fig. 539.

565. Mechanical Effects produced by the Kathode Rays.—Crookes has shown that when the kathode rays fall on a small wheel

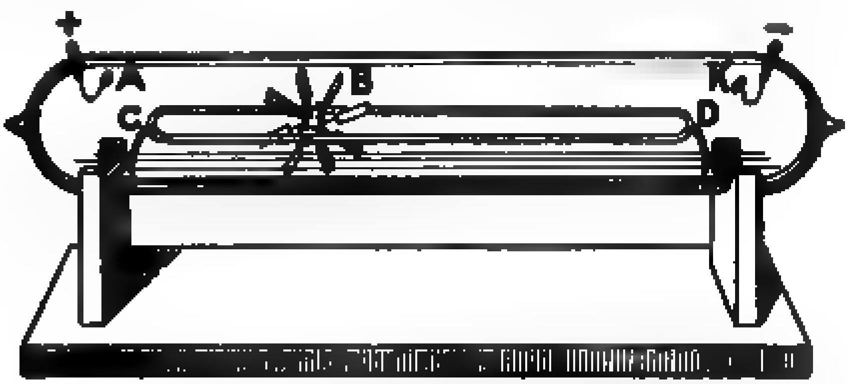


Fig. 540.

with vanes, such as is shown in Fig. 540, in such a way that the rays only strike the vanes on one side of the axis of the wheel this latter will

rotate. By deflecting the rays with a magnet so as to strike the vanes first below the axis and then above, the wheel can be caused to rotate in either direction. It is not certain that this rotation is really due to the momentum of the charged particles of gas shot out from the kathode, and it has been suggested that it is due to secondary thermal effects produced by heat developed in the vanes due to the impact of the kathode rays.

566. Distribution of Potential along an Exhausted Tube during the Passage of a Current.—When a current is passed through a metallic conductor the fall of potential between any two points on the wire is proportional to the current passing and if the wire is uniform the fall of potential along the wire will also be uniform. In the case of an electrolyte, while the fall of potential along the electrolyte is uniform, yet in the cases where there is polarisation at the electrodes an abrupt change of potential takes place at the surface of separation of the liquid and the electrodes. In the case of the passage of electricity through a gas in a rarefied condition in a tube, such as that shown in Fig. 533, the potential gradient, that is, the difference of potential which is shown between two auxiliary electrodes fixed between the kathode and anode divided by the distance between these electrodes, varies greatly from one part of the tube to another. Further, in the positive column the potential gradient is almost independent of the strength of the current passing through the gas. The greater part of the fall of potential which occurs between the anode and the kathode occurs in the immediate neighbourhood of the kathode. Although experiments on this point are subject to considerable question as to what they really represent, they seem to indicate that a considerable potential difference is necessary to cause the charged particles of a gas to give up their charges to a cold metal electrode. If, however, the electrodes are raised to a white heat the passage of electricity from the gas to the electrodes seems to be very much facilitated. Thus in the arc discharge, where the carbons are in air at a very high temperature, the potential difference necessary to continue the discharge through the air is comparatively small compared with what would be necessary to cause a discharge between cold electrodes which are separated by the same distance.

567. Gaseous Dissociation.—It is found that when a spark is passed between two electrodes in air at atmospheric pressure the potential required to start a spark between the electrodes is very much greater than that necessary to continue to force a discharge between the electrodes. This can very clearly be shown by means of the arrangement shown in Fig. 541. A large Leyden jar, A, is connected with a circuit, in which are placed two spark gaps, C and D. A small jar, B, is connected with a circuit, which includes one of these spark gaps, D. If now the jar A is charged up to such a potential that the discharge is not able to pass over

the two spark gaps in its circuit, then on passing a small spark across the gap D, the resistance of this gap will be so much reduced by the passage of this spark that its resistance, together with that of the gap C, is no longer able to support the difference of potential with which the jar A is charged, and so a spark will occur at C and at D due to A.

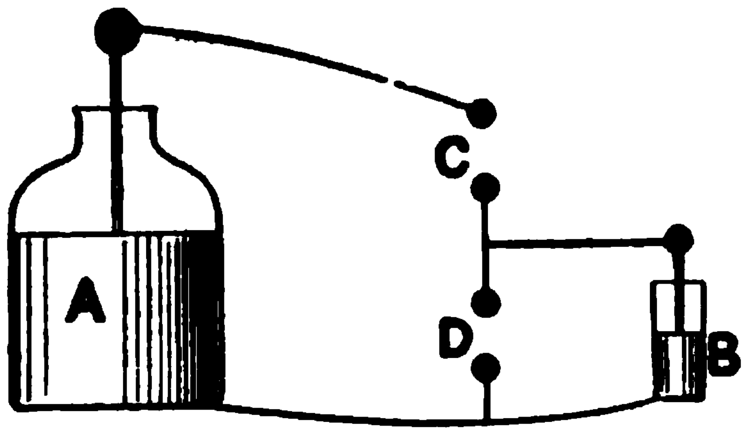


FIG. 541.

This effect has been accounted for by supposing that before an electric discharge can pass through a gas some of the molecules of the gas must be broken up into ions, that is, dissociated. The great initial difference of potential required to cause the passage of the discharge is supposed to be due to the fact that a considerable electromotive intensity is necessary to produce the dissociation of the gas. The dissociation being once produced it is supposed that an appreciable time is required for the dissociated ions either to recombine or to diffuse away from the spark gap, and so the electromotive force necessary to pass a discharge is reduced, since it has not to do the work of dissociation. That there is at any rate some truth in this view is probable from the experimental fact that it is possible to blow out an arc, that is, if fresh undissociated air is blown between the electrodes, the potential difference which was sufficient to maintain a discharge when the dissociated air was allowed to remain between the electrodes will no longer be sufficient to dissociate the fresh air. This view is further supported by some photographs of a spark taken by Feddersen with a rotating mirror, when a current of air was passed between the electrodes.

The first sparks are straight, and stretch straight from one electrode to the other, while the later sparks are blown away in the middle in the direction of the current of air as shown in Fig. 542. It would thus seem that the air first becomes dissociated along the line of maximum electromotive intensity, and this dissociate air being carried away with the air current, the spark prefers to pass along the path of this already dissociated air, although it is longer, than to dissociate a fresh quantity of air. This fact that the air requires a considerable electromotive intensity to cause dissociation, but that once dissociation is produced only a comparatively small difference of potential is required to keep up the discharge, has received a very interesting application at the hands of Hertz, who, as we shall see later, has by this means been able to show experimentally that light really consists of an electro-magnetic disturbance in the ether, as had been predicted by Maxwell.

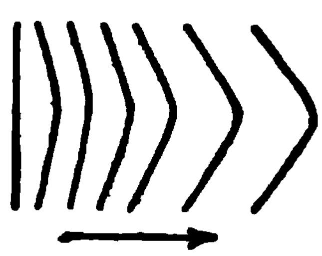


FIG. 542.

A subject which seems to have considerable bearing on this p the passage of electricity through hot gases. When the tempera a gas is raised to that which corresponds to a red heat in a sol electromotive intensity necessary to produce the passage of a disch is very much reduced. The degree with which hot gases conduct tricity is, however, very different for different gases. Thus, such as air, nitrogen, and hydrogen only conduct very feebly even at a temperature, so that a charged body when surrounded by these loses its charge only very slowly, and this loss seems to be due sole a convective action. These gases are those which are known no be easily dissociated when heated. Gases, on the other hand, suc iodine and hydriodic acid, which are known from experiments on density at high temperatures to dissociate into atoms, conduct, w heated to such temperatures, with comparative facility. In the case other gases such as ammonium chloride, which, although they disso when heated, do not split up into atoms but into simpler molecules the case of NH_4Cl into a molecule of ammonia, NH_3, and a molecule hydrochloric acid, HCl), the conductivity when hot is only very sl It would appear from this consideration that the passage of electr through a gas can only occur when there are free atoms present to ca the charge.

The passage of electricity through a gas, even at a low temperatu is very much facilitated if the kathode is exposed to ultra-violet This effect seems to be due to the production of dust in the surround gas by the disintegration of the negatively electrified electrode under influence of the ultra-violet light.

The passage of electricity through a gas is also very much facilitat if the air between the electrodes is traversed by the Röntgen rays. gas which has been exposed to the action of the rays retains its proper of conducting electricity for a considerable time after the rays ha ceased to pass. The gas may also be passed through a considera length of glass-tube without losing its property of conducting. It ap pears, therefore, that the Röntgen rays have the power of ionising th molecules of the gas.

568. Differences between Positive and Negative Electrification.—There are a number of differences between the appearances pre sented by the two electrodes when a discharge passes through a gas, t some of which we have already alluded. Thus in air at atmospher pressure the sparks generally start from one point on the negative electrode, while they spread over a considerable surface of the positive electrode, and in the branched spark all the branches always point towards the negative electrode. In an exhausted tube, as we have seen the appearances at the two electrodes differ in a most marked manner. There are other differences between positive and negative electrifica ture. Thus a piece of bright zinc, when illuminated by ultra-violet light

loses a negative charge, while it is able to retain a positive charge under these conditions. When a brush discharge is formed at a point the potential will be greater if the point is positively electrified than if it is negatively electrified.

The most striking difference is obtained if a discharge is produced between a charged conductor and the surface of a non-conductor on the surface of which some badly conducting powder, such as lycopodium, has been strewn, or on to the surface of a photographic dry-plate. The appearance when the conductor is positively electrified is shown in Fig. 543, while in Fig. 544 the appearance when the conductor is negatively electrified is given, and it will be seen that the difference is most marked. The explanation of these differences has not yet been given, and although there are many other facts which seem to have a bearing on this most fascinating branch of physics, space will not permit of our dealing with them in these pages. The reader who wishes to pursue the subject further will find a very complete account of the work which has been done in this subject in Professor J. J. Thomson's "Recent Researches in Electricity and Magnetism," and also in a volume by him on "The Passage of Electricity through Gases."

FIG. 543.

FIG. 544.

PART IX.—MAXWELL'S ELECTRO-MAGNETIC THEORY

CHAPTER XIX

TRANSFERENCE OF ELECTRO-MAGNETIC ENERGY A MAXWELL'S ELECTRO-MAGNETIC THEORY OF LIGH

569. Poynting's Theory.—We have seen how, according to the of Faraday and Maxwell, if F is the strength of the electrostatic field given point, and K the specific inductive capacity of the medium, energy stored up in each unit of volume of the dielectric at the point is equal to $KF^2/8\pi$. It can also be shown that in the same way energy stored up in each unit of volume of a medium of which the meability is μ at a point of a magnetic field where the strength of the is H, is equal to $\mu H^2/8\pi$. Hence the electric and magnetic energy unit volume of a medium, which is the seat of both electro-static magnetic forces, is $KF^2/8\pi+\mu H^2/8\pi$.

Suppose a condenser, AB, Fig. 545, is charged so that the plate positive, then tubes of force will stretch from the plate A to the plate The greater proportion of these tubes will stretch across the space between the two plates, so that most of the energy due to the charge will be stored up in the dielectric between the plates. The whole energy is in this case the form of electrostatic strain in the

FIG. 545.

medium, since there is no electro-magnetic force produced in the medium. If the plates of the condenser are connected by a conducting wire, which we may suppose of very great resistance, so that the condenser takes an appreciable time to discharge, this wire will be traversed by an electr

current, and at the same time the difference of potential between the plates of the condenser will diminish. During the passage of the electricity through the wire, there will be produced an electro-magnetic field in its neighbourhood, that is, the surrounding medium will possess energy due to the magnetic strain set up. Also the passage of the electricity will be accompanied by the production of heat in the wire, according to Joule's law. When the discharge is complete, the whole of the energy which was originally stored up as electro-static strain of the medium between the plates of the condenser will have been converted into heat in the connecting wire. During the process, however, a certain proportion will have existed in the medium surrounding the wire in the form of energy of the magnetic field, although it also finally becomes changed into heat in the wire. An interesting question now arises as to the way in which the energy travels from the medium between the plates to the wire. Poynting has shown that the energy travels through the medium separating the plates and surrounding the wire, and that the paths along which the energy moves are the intersection of the equipotential surfaces of the electro-static and the electro-magnetic fields. Thus in the case of the condenser discharging through the wire, the tubes of force are supposed to spread out from the space between the plates, the ends of the tubes remaining on the plates. These tubes will meet the wire, and when they do this, they will be broken up and the energy which each contained will be delivered to the wire, where it will appear as heat. The breaking up, or rather absorption, of each tube by the wire will allow another tube to expand from the space between the plates. For each tube, since it exerts a lateral compression on the inside tubes, will tend to prevent their leaving the space between the plates. The absorption of a tube by the wire will reduce this lateral pressure exerted on the inside tubes, and hence more tubes will be able to swell out from the space between the plates.

On Poynting's theory the energy which is transmitted, say, along a telegraph cable is not transmitted along the conducting wire but through the insulating sheath, the object of the wire being to direct the path along which the energy travels.

The telegraph cable may be regarded as a wire surrounded by a concentric conductor, the sheath, the interspace being filled with a dielectric. When the wire is positively electrified and the sheath negatively by connecting the wire, say, to the positive plate of a charged condenser, the negative plate being put to earth, that is, connected to the sheath, tubes of force will stretch across from the wire to the sheath. These tubes will travel forward, each carrying its share of electrical energy. If we suppose the thickness of the insulating covering to remain the same throughout, then the length of the tubes of force will remain the same as they travel onward. The difference of potential between the ends of each tube will, however, diminish as the tube advances, according to

Ohm's law. Since the difference of potential between the wire and the sheath decreases, the work which would have to be done to carry unit charge from the neighbourhood of the wire to that of the sheath will decrease, and since the distance between the two is supposed to remain the same, it must follow that the force acting on the unit charge will also diminish, that is, the strength of the field, F, between the wire and the sheath will decrease as we go from the sending end of the cable. Now we have seen that the energy contained in unit length of each tube of force is equal to $F/2$. Hence, since the length of the tubes remains constant and F decreases, the quantity of energy contained in each tube will decrease as the tube travels away from its starting-point. The passage of the current through the wire and sheath is, we know, accompanied by the conversion of a certain proportion of electrical energy into heat, and this decrease in the electrical energy of each tube as it travels along represents the loss of energy in the conductor, according to Joule's law.

Since the electro-static lines of force are radial, the electro-static equipotential surfaces will be cylinders which are concentric with the wire and the sheath. The magnetic lines of force are circles with the wire as centre and in planes at right angles to the length of the wire, so that the magnetic equipotential surfaces are planes which pass through the wire. The intersection of the two sets of equipotential surfaces will be lines which are parallel to the axis of the wire, and it is along these lines that the energy travels out from the battery at the sending station to the distant end of the telegraph cable.

The supposition that the electro-static equipotential tubes are cylinders of which the axis of the wire is the axis is not quite true, for as we go away from the sending-point the difference of potential between the wire and the sheath will decrease by Ohm's law, so that the number of equipotential surfaces included between the wire and the sheath must decrease, the surfaces being supposed to be drawn for a given difference of potential between consecutive surfaces. The result is that the equipotential surfaces are really frustra of cones. These cones will intersect the wire and the sheath at intervals along the cable, and it is along the line of intersection of such a cone with the magnetic equipotential surfaces that the electrical energy travels which enters the wire or sheath and is converted into heat. If the wire and sheath were composed of conductors of zero resistance there would be no fall of potential along the wire, and in this case the electro-static equipotential surfaces would nowhere intersect either the wire or the sheath, so that no electrical energy would travel into the wire or sheath, and hence no heat would be generated.

When a current is flowing in a circuit, say a coil, the space surrounding the coil will be a magnetic field, and hence there will be a certain amount of energy stored up in this magnetic field. If now the current

is stopped the magnetic field will cease to exist, and the question arises, what becomes of the energy which was stored up in the field? This energy, if the circuit is at a distance from other circuits, returns to the circuit and gives rise to the induced current within the circuit which is produced when the current is stopped. Thus the phenomenon of self-induction (§ 518) is due to the return to the circuit of the energy which during the passage of the current is stored up in the magnetic field produced by the current. When a current is started in a circuit, some of the energy of the battery employed to send the current is used up in providing the energy of the magnetic field. When a second circuit is near the circuit in which the current is flowing, on stopping this current some of the energy of the magnetic field will soak into this neighbouring circuit and will produce in it an induced current.

570. Magnetic Force caused by the Motion of Electro-static Tubes of Force.—We have seen that when electricity moves from one part of a conductor to another, that is, when a current passes through a conductor, that a magnetic field is produced in the neighbourhood of the conductor in which the electricity is moving. It might be conceived that the magnetic field produced in this way by the movement of electricity was due to some special property of the electricity when it is moving from one part of a conductor to another. When a conductor is charged with electricity, the electricity being at rest, the space surrounding the charged body is in such a condition that electro-static forces are set up, that is, it is an electro-static field. In the last section we have seen how the passage of an electric current through a wire, which for simplicity we took double so that the outgoing and return were close together (it must be remembered that there must always be a return; it may be at a considerable distance from the portion of the circuit we are immediately considering, but it is there nevertheless), is accompanied by the motion of the electro-static tubes of force through the medium between the wires. Since by the motion of electricity in a conductor which is accompanied by the motion of the tubes of electro-static force, or, as we may call them, the Faraday tubes, to distinguish them from the tubes of magnetic force, magnetic forces are set up in the dielectric surrounding the conductor, the question arises, would magnetic forces be set up in the same way in the surrounding dielectric if we were to cause the motion of Faraday tubes through the dielectric by moving the body on which a charge exists? Thus suppose we consider two metal plates placed parallel to one another in air, one charged positively and the other negatively. The Faraday tubes will then stretch across from the positive plate to the negative plate, and in the space between the plates we have an electro-static field; but as long as the charges on the plates are at rest there will be no magnetic field. Suppose now the two plates are moved parallel to their own plane and at the same speed, then the tubes will not move with reference to the charged plates but they will sweep

through the air which, owing to the motion of the plates, will pass through the space between the plates. In this case, then, we have produced a motion of the Faraday tubes with reference to the dielectric (air) without any motion of electricity on conductors, and the question arises, will the air between the plates in which the Faraday tubes are moving be the seat of a magnetic field, as it certainly would be if the motion of the tubes were going on owing to the motion of electricity? This question was answered by Rowland, who found by experiment that a magnetic field was produced by the motion of the tubes caused by the motion of the charged body with reference to the dielectric. Since this experiment shows that magnetic force can be produced by the motion of Faraday tubes through a medium, it seems only legitimate to suppose that in every case the production of a magnetic field by a current is due to the motion of the Faraday tubes, which is always going on when such a magnetic field exists. If we adopt Ampère's hypothesis that the magnetism of permanent magnets is due to currents which circulate in the molecules of the iron, then, since these currents must be accompanied by the motion of Faraday tubes, in this case also the magnetic field produced can be considered as due to the motion of these tubes.

Although to go into this subject any further would lead us beyond the scope of this work, we may mention that Professor J. J. Thomson has shown how the various phenomena of the magnetic field can be explained, if we suppose that the motion of the Faraday tubes produces a magnetic field the direction of which is perpendicular both to the length of the tube and to the direction in which the tube is moving.

571. Displacement Currents.—When an electromotive force or difference of potential acts between two points of a conductor, then a motion of electricity is produced in the conductor, that is to say, a current is produced. If the conductor is an electrolyte, then a current is produced, but while there will be a certain amount of energy converted into heat, at the same time the passage of the current will be accompanied by certain chemical changes. In both of these cases the passage of the current will, if we leave out of consideration opposing E.M.F.'s produced by polarisation, continue as long as the difference of potential is maintained. In the case of a difference of potential being produced between two points in a dielectric the circumstances are quite different, for in this case no current passes through the dielectric, neither does any chemical change take place.

As we have seen, however, the dielectric is evidently in a state of strain, for it has become doubly refracting, while the fact that if the difference of potential exceeds a certain value a spark passes, shows that the medium cannot support an indefinitely great electric stress.

In order to account for the different properties of dielectrics, Maxwell supposed that when a dielectric is subjected to an electromotive force, that is, to an electrical stress, a displacement of electricity takes place in

the dielectric in the direction of the electrical stress, but that, unlike what is the case with conductors, a dielectric is able to continuously support the stress, the corresponding strain being the displacement of positive electricity in the direction of the electromotive force and negative electricity in the opposite direction; the difference between a conductor and a dielectric with reference to the electric stress being similar to that between a liquid and a solid with reference to a sheering stress. In the solid a sheering strain is accompanied by a stress which opposes the strain, and is equal and opposite to the strain. In the case of a liquid, however, as long as the strain is changing there will be a certain stress called into play, as we saw when considering the viscosity of liquids, but no permanent strain can be kept up, so that this case corresponds to that of a conductor in which no permanent electrical displacement can be kept up.

Let A and B be two parallel metal plates forming a condenser. Consider any one of the tubes of force stretching between the plates, then this tube will start from a portion of the plate A, containing a unit positive charge, and will end on a portion of the plate B, containing a unit negative charge. Now Maxwell supposes that the charges which appear on the metal plates are simply the manifestations of the state of strain existing in the dielectric contained within the tube of force.

Suppose that A and B (Fig. 546) are the two plates of a condenser, and that these are connected by a wire, W, in which is placed a source of E.M.F., say an electric battery, E. Owing to the action of the E.M.F., suppose that a quantity of electricity, Q, is displaced along the wire, so that A becomes positively electrified and B negatively electrified. The effect of these electrifications of A and B will be to produce an electro-static force acting from A to B in the dielectric between the plates. Now Maxwell supposes that this force will produce an electric displacement within the dielectric, the positive electricity being moved from A towards B, and further, that if we consider any plane, CD, drawn so as to separate the dielectric between the plates into two strata, then the total quantity of electricity which will cross this plane owing to the displacement will be equal to Q. Hence at the same time that, owing to the E.M.F. E, a quantity of electricity, Q, is forced across every cross section of the wire, W, an equal quantity of electricity will be forced across every cross-section of the dielectric. If the E.M.F. in the wire is removed the condenser will discharge through the wire W, and while Q units of electricity will flow through the wire from A to B, Q units will also cross every plane, such as CD, drawn across the dielectric in the direction from B to A. Hence the charge and the discharge of a condenser may be regarded as the motion of electricity in a *closed circuit*, just as is the case when we have to do with

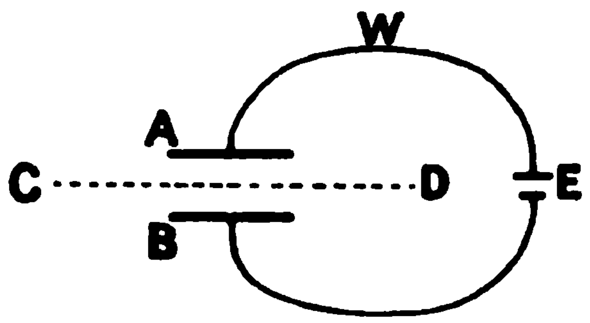


FIG. 546.

circuits which are composed exclusively of conductors. The difference of the two cases lies in the fact that when the circuit consists of conductors only, the resistance which opposes the displacement is independent of the quantity of electricity which has been displaced through the circuit, while, when a portion of the circuit consists of a dielectric, the displacement of electricity through the dielectric is accompanied by the manifestation of a force which opposes the displacement, and which is proportional to the displacement and opposes it. The movement of electricity will therefore go on in the circuit till the opposing force produced in the dielectric, due to the displacement, will be exactly equal and opposite to the E.M.F. which tends to produce a current in the circuit. On the removal of this E.M.F., then, just as a deformed spring on the removal of the deforming force will spring back to its original shape, so the dielectric elastic force, being no longer opposed by the external applied deforming force, will cause the springing back of the electrical displacement, that is, will cause the passage of a current in the circuit in the opposite direction to the charging current. According to Maxwell's theory of electricity, in considering the magnetic actions which accompany the charge and discharge of the condenser, we must include not only the actions of the currents which flow in the wire W, but also the magnetic effects of the displacements which take place in the dielectric. In other words, magnetic effects may be produced both by conduction currents, such as we have exclusively considered hitherto, and also by displacement currents in the dielectric.

Let us now consider more in detail the effects of regarding a dielectric as an elastic medium in which a stress, due to the action of electro-static force, or rather induction, produces electrical displacement. Let us consider the case of two infinite plane conducting plates placed parallel to one another at a distance d apart. If σ is the charge on each unit of area of the positively charged plate, then there will be σ tubes of force starting from each unit of area of this plate. Since the plates are infinite, the tubes of force are of uniform cross-section, and are everywhere at right angles to the plates. If we consider a single tube of force, its cross-section will be $1/\sigma$. The total displacement across any plane drawn between the plates will be equal to the charge on the plates, and since the plates are so large that the effects of the edge may be neglected, the displacement through each unit of area of a plane drawn parallel to the plates will, at any rate near the centre, be equal to σ. Now, if the specific inductive capacity of the dielectric is K, the force acting at any point between the plates will be $4\pi\sigma/K$. Hence the stress acting to produce electrical displacement is $4\pi\sigma/K$, and the strain produced across unit area, that is, the displacement across unit area, is σ. Now, when considering the effects of a stress on an elastic material (§ 122), the ratio of the stress to the strain was called the elasticity of the material. Hence by analogy we may call the ratio of the electric force to the displacement

it produces the coefficient of elasticity of the medium. Thus the electrical coefficient of elasticity of a dielectric is equal to $4\pi/K$, where K is the specific inductive capacity of the medium.

The reason why, although displacement takes place throughout the mass of the dielectric, it is only on the bounding surface between the dielectric and a conductor that an electrical charge is manifest, is similar to the reason that although a magnetised bar is magnetised throughout its mass, it is only at the ends that this magnetism is evident. Within the mass of the magnetic material the opposite magnetisation of adjacent portions of the material neutralise each other's external effect. In the same way, in the case of the dielectric, although there is displacement throughout the mass, yet, since the displacement in any small portion of the medium will cause it to become electrified positively on one side and negatively on the other, these charges will not produce any external effect, since the neighbouring portions of the medium will also be electrified in such a way that their sides turned towards the first portion will exhibit an equal and opposite electrification.

572. Maxwell's Electro-magnetic Theory of Light.—Although Maxwell's theory of electrical displacement does not in any way attempt to tell us what electricity is, yet, by showing how the observed facts can be accounted for by ascribing certain elastic properties to the medium, it is of very great importance, and it led Maxwell himself to the important conclusion that it must be possible to produce waves in a dielectric, the periodic disturbance by which they are constituted being electric displacement currents in the dielectric. Further, an examination of the properties of such waves showed that they will be propagated with a definite velocity, this velocity being that with which light is propagated in the given medium. The medium in which the waves are propagated is not matter, for electrical forces can be transmitted through a vacuum, so that we are led to postulate on this account the existence of an ether which pervades all space.

Although the matter is not the medium in which the waves are produced, there is no doubt that the presence of matter does influence the velocity with which the electrical waves are propagated. Now when considering the propagation of light we have been led to similar conclusions, for the velocity of light depends on the nature of the matter occupying the space through which the light is travelling, and since light can travel through space where, as far as we can tell, no matter exists, some other medium besides ordinary matter has to be postulated. Hence what, till Maxwell's time, were regarded as two entirely distinct sciences, namely, light and electricity, lead to the postulation of the existence of an ether, and since the velocity with which waves of electrical disturbance travel through the ether was found by Maxwell, according to his way of regarding the phenomena, to be the same as the velocity of light, he naturally concluded that the two phenomena were identical, and that

that which we call light is really due to the passage of electrical waves. In some such way was Maxwell led to his electro-magnetic theory of light, which, when it was first proposed, received hardly any support, but which now is accepted by every one, and we shall now proceed to review very briefly some of the experimental evidence in support of this theory.

If we consider the vibrations set up in a spring which is clamped at one end, it is evident that two conditions have to be fulfilled for these to take place. In the first place, there must be an elastic force called into play when the spring is bent, tending to move the spring back into its undeflected position. This alone is not, however, sufficient, for unless the spring possessed inertia it would not possess any kinetic energy when it reached its position of equilibrium, and therefore would not swing past this position against the elastic forces which oppose its motion. In the electrical case we have just seen how a consideration of the electrical properties of dielectrics has led us to endow them with electrical elasticity, and we have now to consider what evidence there is for supposing that electricity possesses something of the nature of inertia, for if electrical waves are to occur it must possess such inertia. In § 518 we saw how, when we attempt to start or stop a current in a circuit, or more generally, if by any means we attempt to alter the number of magnetic tubes of induction passing through a circuit, electrical forces are called into play which tend to oppose such an alteration. Thus the phenomena of self and mutual induction both indicate that electricity possesses something of the nature of momentum.

In the case of a moving body the momentum is equal to the product of the mass into the velocity, that is, into the rate of change of the position of the body. The force which produces the momentum is equal to the rate of change of the momentum, that is, the product of the mass into the acceleration or the change of position per second per second.

In the electrical case the electromotive force necessary to change a current flowing in a circuit is equal to the rate of change of the induction, B, through the circuit. The induction through a circuit is equal to the product of the magnetising force into the permeability, μ, while the magnetising force is proportional to the rate of change of the electrical displacement in the circuit. Hence the electrical force necessary to change the induction through a given circuit is proportional to the product of the permeability into the change of displacement per second per second. Comparing this case with the mechanical case considered above, we see that permeability in the electrical case plays the same part as mass does in the mechanical one.

The velocity with which an undulatory disturbance will be propagated through a medium being equal to the square root of the quotient of the elasticity of the medium by the density, it will be understood, from what has been said above, how Maxwell's expression for the velocity, v, of electrical waves, $v = \sqrt{K/\mu}$, is obtained.

We have seen that the value of $\sqrt{K/\mu}$, as obtained from a comparison of the values of the electro-static units with the corresponding electro-magnetic ones, was equal to the velocity of light, so that this formula of Maxwell's shows that electrical waves will travel with the velocity of light.

When electrical waves are passing through a dielectric, then, at any point we shall have an electrical displacement produced which will be in a direction at right angles to the direction of motion of the waves. The displacement will occur first in one direction, reach a maximum value, gradually decrease to zero, and then become negative, and so on. Thus the electrical displacement will play the same part in electrical waves as does the displacement in a vertical direction of the water particles in a water wave. As we have seen, the displacement within a dielectric is accompanied by a stress which opposes the displacement, and this stress plays the same part as the action of gravity in the case of water waves.

Suppose we consider a cylindrical portion, AB (Fig. 547), of a medium through which electrical waves are passing, the direction in which the waves are moving being at right angles to the axis of the cylinder and as shown by the arrow. As the waves pass, the electrical displacements in the cylinder AB will take place parallel to the axis, that is, at right angles to the direction of motion of the waves. The sense of the displacement will be alternately in the direction AB and in the direction BA. Now a displacement in the direction AB will produce the same magnetic field as a current in the cylinder from A to B, and will therefore produce a system of magnetic lines of force which will be a series of circles having their centres on AB and lying in planes at right angles to AB. Hence the wave of electro-static displacement will be accompanied by a wave of magnetic force, for when the displacement changes sign the direction of the magnetic force will also change sign. From considerations similar to those adopted in § 273, when considering Huyghens's construction for the wave-front, it will be evident that the only portion of the line of force due to the cylinder AB that will produce a magnetic field will be that portion which is perpendicular to the direction of motion of the wave, that is, the portion in the wave-front. For if we imagine a second cylinder in the dielectric alongside AB, then, if these are both in the wave-front, the displacement currents in them will be in the same phase, and hence the lines of magnetic force in the space between them will be in the opposite direction, and will therefore interfere with one another. Hence every electrical wave will be accompanied by a magnetic wave, the directions of the electrical displacement and the magnetic force being at right angles, but both being in the wave-front. Since it is impossible to obtain one wave without the other, we shall often speak of the one

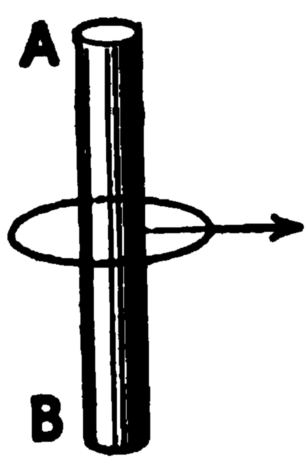


FIG. 547.

only when discussing the phenomena of electro-magnetic waves; it must be remembered that the other always exists.

A somewhat more concrete picture of the condition of a dielectric through which electric waves are passing may be formed by considering the motion of Faraday tubes.

Along each Faraday tube there exists an electrical displacement, and hence, when a tube moves through a dielectric, the portion of the dielectric which at any given instant is included within the tube is the seat of an electrical displacement. The displacement takes place in the direction of the length of the tube and towards the positive end.[1] Thus if we have a series of tubes, such as those shown in Fig. 548, moving in the direction of the arrow, the displacement produced at any point within the dielectric will be upwards when any of the tubes which have their positive ends upwards are passing the point, and downwards whenever one of the tubes having its negative end upwards are passing.

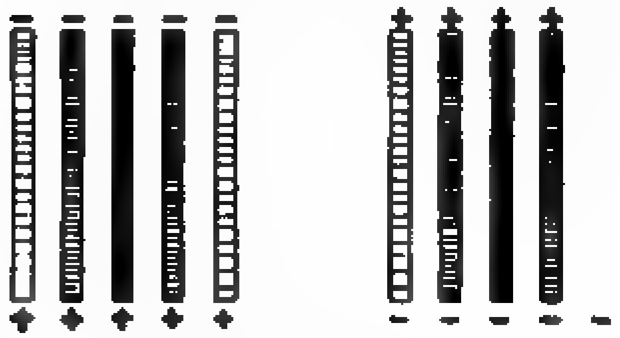


FIG. 548.

In this manner of picturing the passage of electrical waves, the accompanying magnetic field is that which we have already seen occurs whenever we have motion of Faraday tubes, the direction of the magnetic field being at right angles both to the length of the tubes and to the direction of their motion, that is, at right angles to the plane of the paper.

Since each Faraday tube is the seat of a certain amount of energy stored up in the form of electrical strain, this energy will be carried forward by the motion of the tubes, and so we have here a picture of how the energy corresponding to the waves travels. Each tube behaves very much like a stretched rod of india-rubber, for such a rod would possess energy owing to its strained condition, and would be made to do work while regaining its unstrained condition. There is, however, this important difference, that in the case of the rubber the portion of matter which is in the state of strain is carried forward. In the electrical case it is otherwise, for the strain in the ether is handed on from one portion to the next, and at present the mechanism by which this handing on is performed, as well as the nature of the electrical strain

[1] Of course, by the term direction of the displacement, we refer to the direction of the displacement of positive electricity. There will be a displacement of negative electricity in the opposite direction, but as the displacement of positive electricity in one direction is equivalent to the displacement of negative electricity in the opposite direction, we need only speak of the displacement of the positive electricity.

itself, is unknown, and till these are known we are unable to answer the question, "What is electricity?" Since the motion of the energy takes place at right angles to the tubes of force, that is, to the direction of the electro-static field, and also at right angles to the magnetic field, we have here a confirmation of Poynting's theory on this subject (§ 569).

573. Connection between Refractive Index and Specific Inductive Capacity.—If v is the velocity of electro-magnetic waves in air, then, according to Maxwell's theory, we have $v = \sqrt{K/\mu}$, where μ and K are the permeability and specific inductive capacity of air. Similarly, if v' is the velocity in a medium for which the permeability and specific inductive capacity are μ' and K', then $v' = \sqrt{K'/\mu'}$.

Thus
$$v'/v = \sqrt{(K\mu'/K'\mu)}.$$

Now in the case of all transparent bodies μ is very nearly unity, so that in this case we have
$$v'/v = \sqrt{K'/K}.$$

But the ratio of the velocity of light in air to the velocity in a given medium is called the refractive index of the medium, while the ratio K'/K is the specific inductive capacity of the medium taken with reference to air. Thus if n is the refractive index, and K the specific inductive capacity, both taken with reference to air, we have
$$n = \sqrt{K}.$$

That is, the refractive index is equal to the square root of the specific inductive capacity.

When we attempt to test the accuracy of this conclusion by experiment, we are met with the difficulty that since the refractive index changes with the wave-length, that is, the velocity changes with the wave-length of the light, the question arises, what wave-length are we to employ? It is evident that the correct wave-length will be that which corresponds to experiments made when determining K. Now measurements made of the specific inductive capacity by means of the ordinary methods with condensers, are made with alternating currents to avoid the effect of absorption, but the alternations have a frequency of, at most, a few thousands per second. Hence the refractive index which has to be used in testing Maxwell's formula is that which corresponds to a very small frequency, that is, to a very long wave-length; in fact, the wave-length of a light wave of which the frequency is a thousand would be 3×10^7 cm. Now measurements of refractive index can only be made for comparatively short wave-lengths, and it is only by exterpolation that we can calculate what the refractive index would be for very great wave-lengths, and most of the differences in the annexed table are probably due to this cause, for we have no evidence that the laws of the change of refractive index with wave-length derived from the small

range of wave-lengths over which we are able to make experiments will hold over very much greater ranges.

Substance.	Specific Inductive Capacity.	Square of Refractive Index.
Paraffin	2.3	2.02
Petroleum	2.07	2.08
Turpentine	2.23	2.13
Ozokerite	2.13	2.09
Olive oil	3.16	2.13
Benzine	2.22	2.24*
Toluene	2.30	2.25*
Carbon bisulphide	2.67	2.67*
Water	76	1.78*
Alcohol	26.5	1.83*

* These values of the square of the refractive index are for *D*-light.

It will be seen that in general the agreement is satisfactory. In some cases, such as water and alcohol, however, the values obtained for the specific inductive capacity are very much greater than Maxwell's theory would indicate. In the case of water, it has been found that the refractive index for electrical waves having a frequency of about 50×10^6 is 8·9. Hence for waves of this frequency the square of the refractive index, 79.2, is equal to the specific inductive capacity.

574. Transmission of Light and Conductivity.—Electrical waves can only be transmitted through a medium in which an electrical displacement calls forth an elastic resistance, for otherwise a vibratory motion is impossible. In a conductor of electricity, however, electrical displacement can take place, and no force will be called into play tending to oppose the displacement. Electrical waves cannot, therefore, be transmitted through a conducting medium, and since light waves are also electromagnetic waves, they also will not be transmitted through a conducting medium. Maxwell's theory thus explains why the metals are, without exception, opaque to light. Insulators or dielectrics, on the other hand, since they can transmit electrical waves, will also transmit light. It does not follow that if a body will not transmit light that it must be a conductor, for a medium may be opaque because its structure is not homogeneous. Thus glass in a block is transparent, but pounded glass is opaque, the opacity being due to the scattering of the light by the small particles of glass, since there will be a certain amount of reflection at every surface.

575. The Faraday Effect.—In 1845 Faraday discovered that when a beam of plane polarised light (§ 400) is passed through a magnetic field in the direction of the lines of force, the plane of polarisation of the light is rotated owing to its passage through the field. Thus if the light from the source L (Fig. 549) is passed through a polarising Nicol, P, then

through a tube T containing water, or better, carbon bisulphide, and finally through an analysing Nicol A, then, on rotating this analyser so that its principal plane is perpendicular to that of the polarising Nicol, no light will be transmitted. If, however, a current is passed through a coil C which surrounds the tube T, so as to produce a magnetic field with the lines of force parallel to the direction in which the light is travelling, the light will be found to pass through the analyser A. By turning the analyser it is, however, possible to again cut off all the light. This experiment, therefore, shows not only that the plane of polarisation of the light has been rotated, but also, since by rotating the analyser it is possible to cut off all the light, that the beam must remain plane polarised. If the direction of the current is reversed, the direction of the rotation is also reversed.

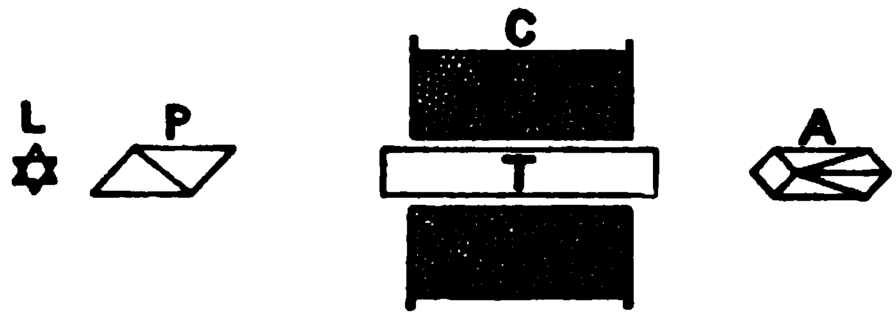


FIG. 549.

There is an important difference between the rotation of the plane of polarisation thus produced by matter when placed in a magnetic field and that produced when a ray of light is transmitted through a plate of an ælotropic body such as quartz (§ 411). Suppose a ray of plane polarised light is transmitted through a tube containing water, T (Fig. 549), in the same direction as that in which the lines of force of the field proceed. Then, looking in the direction in which the lines of force run, the plane of polarisation will be rotated in the clockwise direction. If the direction of the light is reversed, the rotation will still take place in the clockwise direction, as seen by an observer looking along the direction of the lines of force, but will appear in the opposite direction to an observer looking in the direction in which the light is travelling. Hence, if the ray of light, after having once passed through the tube of water in the magnetic field, is reflected back along its course, it will be again rotated in the same direction, as far as the coil is concerned, as during its first passage, and the plane of polarisation will therefore be turned through twice the angle through which it was turned owing to the single passage. In the case, however, of a ray of plane polarised light transmitted through a plate of quartz, Q (Fig. 550), in a direction parallel to the axis of the crystal, the rotation will take place in one direction when the light passes one way, but will take place in the opposite direction, as far as the crystal is concerned, if the direction of the light is reversed. Hence, if a ray of plane polarised light is transmitted through such a plate of quartz, and is then reflected so as to again traverse the crystal in the reverse direction,

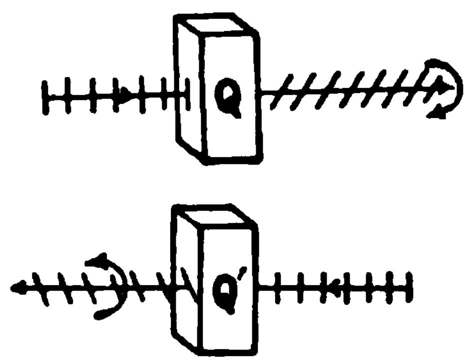


FIG. 550.

the rotation during the second passage will be opposite to that during the first, and on the whole the plane of polarisation will not be rotated.

576. Verdet's Constant.—Verdet, who made a large number of measurements on the magnetic rotation of the plane of polarisation in different media, found that the rotation for any one medium obeyed the following law. If the length of the medium traversed by the light in the direction of the lines of force of a uniform magnetic field of strength H is L, the rotation ϕ produced is given by

$$\phi=\gamma LH,$$

where γ is a constant dependent on the nature of the medium and the wave-length of the light employed, and is called Verdet's constant. The value of γ is taken to be positive when the direction of rotation is the same as that of the current in the coil producing the magnetic field. If L and H are each unity, then the rotation produced is equal to the value of Verdet's constant, so that we may define γ as the rotation produced by unit length of the given substance when placed in a magnetic field of unit strength. It is usual to measure the rotation in minutes of arc, so that in the following table the values given represent the rotations in minutes produced by 1 cm. of the substance in a field of which the strength is 1 *c.g.s.* unit. The light for which the values of γ are given is yellow sodium light.

VERDET'S CONSTANT FOR *D*-LIGHT.

Substance.	Temperature.	Verdet's Constant.
Water	0° C.	0.0131
Carbon bisulphide . .	0° C.	0.0435
Benzine	20° C.	0.0297
Glass (dense flint) . .	15° C.	0.06

The value of Verdet's constant decreases with increase of temperature. The change in most cases is proportional to the change in temperature, water, however, being an exception.

In the case of water and carbon bisulphide, the value of Verdet's constant at a temperature t is given by the following expressions :—

Carbon bisulphide . . $\gamma_t=0.04347\ (1-0.001696t)$
Water . . . . . . $\gamma_t=0.01311\ (1-0.0000305t-0.00000305t^2)$.

In the case of solutions of ferric chloride in water, the rotation is in the negative direction.

When polarised light is transmitted through very thin films of the magnetic metals, iron, nickel, and cobalt, placed in a magnetic field, the plane of polarisation is rotated. In this case, however, the quotient ϕ/LH is not constant, but depends on the value of the magnetic field. H. du Bois has shown that although in the case of magnetic metals

Verdet's constant varies with the magnetic field, if the value is divided by the susceptibility, then the quotient is constant.

577. The Kerr Phenomenon.—Another effect of magnetism on light has been discovered by Kerr, who found that if plane polarised light is reflected from the polished pole of a strong magnet the plane of polarisation is rotated. The direction of rotation when the light is reflected from a north pole is in the clockwise direction, that is, in the opposite direction to that in which a current would have to flow round a coil so as to produce the magnetisation of the magnet.

578. The Zeeman Effect.—In 1897 Zeeman discovered another connection between magnetism and light. He found that if a flame coloured with common salt is placed between the poles of a powerful electro-magnet, and the light given by the flame is examined with a spectroscope of great dispersive power, the appearance of the *D*-lines is greatly altered. If the source of light is viewed at right angles to the lines of force of the field, then recent examination with very powerful magnetic fields and great dispersion has shown that D_1 becomes converted into four lines, while D_2 becomes a sextet. In each case the two central lines are plane polarised, the vibrations taking place at right angles to the length of the line. The outer lines are also plane polarised, but the vibrations are in a direction parallel to the length of the lines. A more usual type of line is one in which a single line becomes, when viewed at right angles to the magnetic field, transformed into a triplet, in which the vibrations in the central line take place at right angles to the length, and in the side lines parallel to the length of the line.

If the source of light is viewed in the direction of the lines of force, the outer components of the triplet obtained are circularly polarised in opposite directions, while the central line is plane polarised.

Lorentz and Larmor have shown that the Zeeman effect can be accounted for if we assume that in all bodies there are present small electrically charged particles which have a definite mass, and that all electrical phenomena are due to the configuration and motion of these charged particles or electrons, while light is produced by the vibration of these electrons. When these electrons move in a magnetic field their natural periods will be subjected to perturbations, owing to the action of the field, and these perturbations will be such as would account for the differences in period indicated by the duplicated lines obtained. From the amount of the change in period produced by a given field, it is possible to calculate the ratio of the charge on each electron to its mass. In § 563 we have mentioned that Professor J. J. Thomson had calculated the mass of the electric carriers in the kathode rays, and it is interesting to note that, if we suppose that these carriers are simply electrons, then the masses, as calculated from the Zeeman effect and the kathode rays, agree. On this hypothesis the molecule of a gas would consist of about 1000 electrons.

CHAPTER XX

ELECTRICAL OSCILLATIONS

579. Oscillatory Discharge of a Leyden Jar. — When a condenser, such as a Leyden jar, is charged, say the inside coating being at the higher potential, there will be a displacement in the dielectric separating the coatings. When the jar is discharged by connecting its coatings by a conducting wire, the displacement decreases till it becomes zero, but when this point is reached under certain circumstances, which we shall consider later, the inertia of the electrical displacement carries it through its position of equilibrium, and a displacement in the opposite direction to the original one occurs. This displacement corresponds to the charging of the jar in the opposite direction, that is, the inside coating becomes negatively charged. As this charging in the reverse direction proceeds, that is, as the negative displacement increases, an opposing elastic force will be called into play which will diminish the electrical kinetic energy till, when the whole of this energy is converted into potential energy in the form of dielectric strain, the jar will start discharging in the opposite direction. The negative displacement will then decrease, becoming zero, and then a displacement will occur in the positive direction, the inside coating again becoming positively charged. Thus the discharge of the jar does not consist of a simple passage of a current in one direction, but the charge surges backwards and forwards, each coating becoming charged alternately positively and negatively, so that an alternating or oscillating current is set up both in the wire connecting the coatings, where the current is a conduction current, and also in the dielectric, where it is a displacement current. The magnitude of the charge decreases with each oscillation, for the passage of the current through the wire is accompanied by the development of heat, according to Joule's law, and this energy has to be supplied by the electrical energy which was originally stored up by the strain of the dielectric. The phenomenon of the oscillatory discharge of a condenser is exactly the same as that of the vibration of a flexible rod clamped at one end. When the free end of the rod is at the extremity of its swing its energy is entirely potential, due to the strain set up. The condition of the rod now corresponds to that of the jar when it has its maximum charge, and possesses energy due to the strain of the dielectric. As the rod swings towards its position of rest, the potential energy becomes gradually con-

verted into kinetic energy; while in the electrical case, as the discharge proceeds, the potential energy becomes converted into kinetic energy, that is, into the energy of the magnetic field produced by the current which flows in the wire and in the dielectric. This kinetic energy carries the rod in the one case, and the electrical system in the other, beyond the position of rest, but now, since the elastic forces oppose the motion, the kinetic energy will gradually be converted into potential energy. When the whole of the kinetic energy is thus converted, the rod will have reached its maximum displacement in the new direction, while the jar will in the same way have its maximum charge in the new direction.

On account of the viscosity of the metal and of the resistance of the air, some of the energy of the vibrating bar will be converted into heat, and the amplitude of the oscillations will decrease. In the same way, in the case of the jar, owing to the resistance of the wire connecting the coatings, electrical energy will be converted into heat, and the amplitude of the electrical oscillations will decrease. The greater the resistance of the wire, the more rapid will be the rate at which the electrical energy will be converted into heat, and the greater will be the rate of decay of the oscillations. If the resistance of the wire is gradually increased, then, just as when a pendulum when displaced in a very viscous material, such as treacle, will not vibrate but will simply slowly move back to its position of rest, so in the electrical case, if the resistance of the connecting wire is very great, no electrical oscillations will be set up, but the charge of the jar will slowly decrease to zero, and will not overshoot this position.

The first to show from mathematical principles that the discharge of a Leyden jar must, so long as the resistance of the discharge wire is not too great, be of an oscillatory character was Lord Kelvin, who also showed that the period of the oscillations (T), that is, the interval between when the jar has its maximum charge in one direction, must be given by the equation

$$T=2\pi\sqrt{LC}$$

where C is the capacity of the condenser, and L is the coefficient of self-induction of the wire connecting the coatings.

The truth of this formula has been tested experimentally by examining the spark which passes between two knobs placed in the discharge circuit of a Leyden jar by means of a rotating mirror. If the discharge consisted of the passage of electricity in one direction only, then when examined with the rotating mirror the spark would appear as a continuous band of light, the length to which it was drawn out depending on the time the spark lasted and on the speed of rotation of the mirror. As a matter of fact, when the resistance of the discharge circuit is not very great, the spark is seen to consist of a number of bright patches separated by dark intervals. Each of the bright patches corresponds to the passage of an

electric current between the knobs, and the fact that the current is not continuous, but consists of a number of separate currents, shows that the discharge must be oscillatory. Further, from the speed of rotation of the mirror and the distance between the images of the separate sparks between the knobs, it is possible to obtain the time of oscillation of the discharge, and the number thus obtained agrees with that calculated from the capacity of the jar and the self-induction of the discharge circuit.

In Fig. 551 a copy of a photograph taken by Boys, of the spark produced between two knobs included in the circuit of a condenser by means of what was the equivalent of a rotating mirror, very clearly indicates the oscillatory nature of the discharge. It will be noticed that the brilliancy of the spark decreases with each oscillation, this being due to the energy spent in the wire and also in the spark gap, where the light produced is evidence of the dissipation of electrical energy. It is rather interesting to remember that, according to the electro-magnetic theory of light, light consists of electrical oscillations; thus at the spark gap

FIG. 551.

we have the comparatively slow electrical oscillations which are taking place in the jar circuit partly converted into the much more rapid electrical oscillations which are capable of affecting our eyes, and which we call light. Much the same thing occurs in the wire, for the heat which is there developed raises the temperature of the wire, and hence it commences sending out radiant heat, which is simply electro-magnetic waves of which the wave-length, while being much smaller than that of the electrical waves in the jar circuit, are yet too long to affect our eyes as light.

580. Resonance in Leyden Jar Circuits.—When dealing with sound, we found that when a tuning-fork is in the neighbourhood of another fork which is in vibration, this latter will send out sound-waves which will strike the other fork, and if the pitch of the forks is the same, these waves, which are incident on the second fork, will set it in vibration. It is, however, only when the two forks are in unison that sounding the one will set the other in vibration. Suppose now we have a Leyden jar in which an oscillatory discharge is taking place; then, according to

Maxwell's theory, electro-magnetic waves will be produced in the medium surrounding the jar which will be of the same frequency as the oscillations in the jar. Hence, if a second jar is placed in the neighbourhood of the first, and its capacity and the self-induction of its discharge circuit are so chosen that the frequency of the electrical oscillations which would occur in it when it is charged and then discharged is the same as that of the oscillations in the other jar, then we should expect that the second jar would respond to the first, and that electrical oscillations would be set up in it by resonance.

The correctness of the above view has been very clearly demonstrated by Lodge, who employed the arrangement shown in Fig. 552. A Leyden jar A has its inside and outside coatings connected by a wire circuit CGD, in which a spark gap G is included. The terminals of an electrical machine are connected to the outside and inside coatings of this jar, so that when the machine is worked the jar will become charged till the difference of potential between the knobs of the spark gap is sufficiently great to force a spark across. Now, as we have seen in § 567, when a spark passes in a gas, the resistance of the gas in the path of the spark becomes very much reduced. The result is that the passage of a spark at G, since it makes the air between the knobs a conductor, has the effect of converting the broken circuit of the jar A into what is practically a complete conducting circuit. Hence the jar discharges, and electrical oscillations are set up. A second jar B, of the same capacity as the first, is placed in the neighbourhood, and is fitted with a conducting circuit EF connecting its inside and outside coatings. A strip of tinfoil connected with the inside coating of the jar B is brought over the edge of the glass to within about a millimetre of the top of the outside coating. The spark gap H thus formed serves as an indicator to show when oscillations are set up in the jar B, for when these occur the inside and outside coatings will, at the extremity, so to speak, of each oscillation, be at different potentials, and so a spark will tend to pass at H. The length of the discharge circuit of B, and hence the self-induction of this circuit, can be altered by sliding the wire EK through a ring at F. Now altering the self-induction of the discharge circuit will alter the frequency of the oscillations in the jar, for the periodic time is equal to $2\pi\sqrt{LC}$,

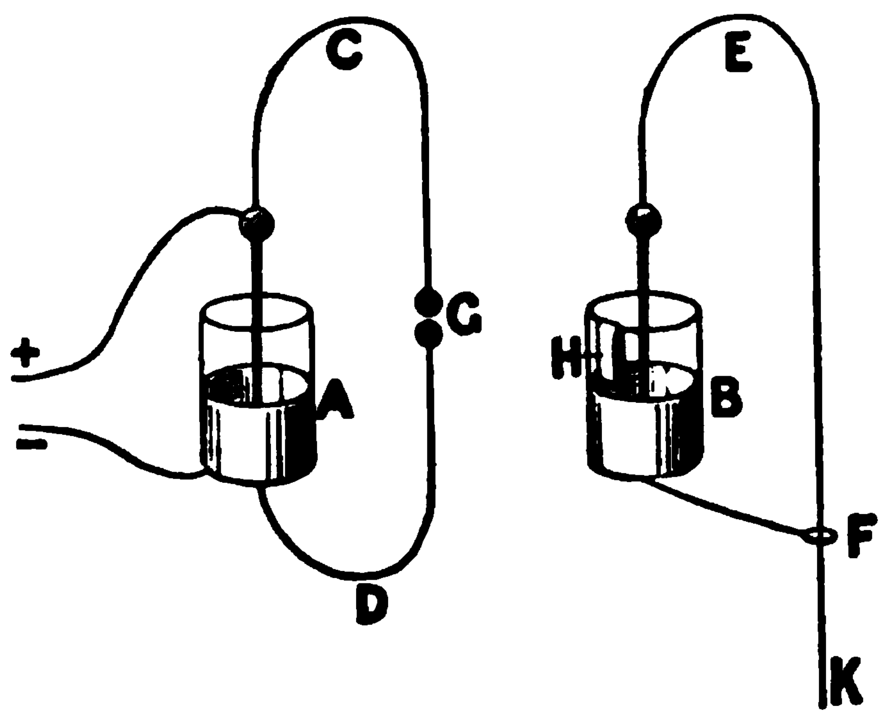


FIG. 552.

where L is the self-induction of the circuit. Hence, by drawing the wire EK through the ring, we can alter the frequency of the electrical oscillations corresponding to the jar B, or in other words we can tune the jar. Now it is found that the jar B responds, that is, induced oscillations as indicated by the sparks at H only occur, when the length of the discharge circuit has a particular value. If the circuit is longer, so that the natural period of the jar B is greater than that of A, there is no response; while if it is shorter, so that the period is less, there is also no response. We have here then a case which is completely analogous to the case of the response of two tuning-forks.

581. Electrical Oscillations of Small Wave-Length.—The periodic time of the electrical oscillations set up when an ordinary Leyden jar is discharged through such a circuit as shown in Fig. 552 is about 1.5×10^{-7}, and since electrical waves travel with the velocity of light, the wave-length, which is the distance through which the disturbance travels during the periodic time, will be about $3 \times 10^{10} \times 1.5 \times 10^{-7}$, or 4.5×10^3 cm. Thus the waves given out by such a jar are of quite unmanageable length. Now the experiment of two jars tuned to unison considered in the last section, although it shows that energy is communicated from one jar across the intervening air to the other, does not indicate whether or not the disturbance travels through the intervening air instantaneously, or whether it takes time to travel from the one to the other, as Maxwell's theory indicates. In order fully to test the question of the propagation of electro-magnetic waves through a dielectric, we require therefore to make some other experiment, and this involves the production of electrical waves of smaller wave-length than those given by the jar. Before proceeding to describe how this was accomplished, it will be worth while to consider for a moment the conditions which have to be fulfilled in order to produce waves of comparatively small wave-length. Taking first the case of a pendulum; in order to set it swinging, we not only have to pull it to one side, but we must let the bob go in a time which is short compared to the periodic time of the pendulum. Suppose that the bob is pulled aside by means of a string, then if, instead of either breaking the string or letting it go, we allow it to run through our fingers, the bob will simply slowly return to its position of rest, and will not be set in vibration. The same considerations apply in the case of attempting to obtain electrical oscillations of small wave-length, only in this case the time of a complete vibration, when the wave-length is a metre, is only 3.3×10^{-9} second. The problem, therefore, is to charge up a condenser, the term condenser being here used in a general sense for any two conductors which are charged simultaneously to opposite electrifications, and of course for this purpose there must be a break in the discharge circuit to allow of the difference of potential between the plates being produced, and then by some means to suddenly close the break in the circuit. Any attempt to perform this

closing by means of any mechanical device, such as a key, would be futile, for the time during which the closing was taking place would be much greater than 10^{-6} second. To Hertz belongs the honour of having discovered a method of overcoming this difficulty, and thus rendering the production of electro-magnetic waves of small wave-length possible. He found that the electric spark, when it passes in air between two knobs which are brightly polished, has the remarkable property of not only making the air between the knobs for the time being a conductor, but it performs the change from the condition of a comparatively perfect insulator, which exists before the passage of the spark, to the comparatively conducting condition, which holds after the passage of the spark, in a time which is small compared even with 10^{-6} second. Thus the spark gap in the discharge circuit performs a twofold duty. It first acts as an insulator, and so allows the conductors on the opposite sides to be charged to an appreciable difference of potential, that is, it allows an appreciable amount of energy to be stored up in the condenser, and then, when the difference of potential has reached a certain value, it suddenly releases the electrical strain, by converting the air between the knobs into a conductor, and so allows the strained dielectric in the remainder of the field to recover, and in doing so to set up a current in the circuit.

582. Hertz's Experiments.—The arrangement employed for producing the electrical oscillations used by Hertz, and called an oscillator, is shown in Fig. 553. The terminals of an induction coil, C, are connected with two metal rods, on each of which a metal sphere, A and B, is threaded. The ends of these rods are supplied with well-polished brass knobs which form the spark gap G. When the coil works, each time that the primary circuit is broken the induced E.M.F. produced in the secondary circuit charges the conductors A and B till the potential difference between the knobs of the spark gap is sufficiently great to cause the passage of a spark. When the spark passes, oscillations are set up between the spheres which gradually die out on account of the energy being partly converted into heat in the rods connecting the spheres and in the spark between the knobs, and partly radiated into the surrounding space as electro-magnetic waves, each wave representing of course a certain amount of energy which has been lost by the oscillator.

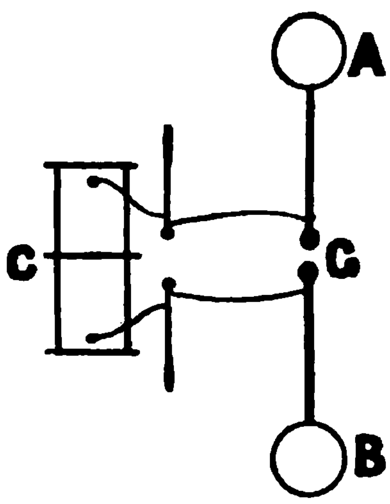


FIG. 553.

In Hertz's original oscillator, in which the diameter of the spheres was 30 cm. and the distance between the centres was 100 cm., the period, as calculated by Lord Kelvin's formula, was 1.85×10^{-8} second, so that the wave-length was $3 \times 10^{10} \times 1.85 \times 10^{-8}$, or 5.55×10^{2} cm. In order to obtain waves of still smaller wave-length, Righi devised the

arrangement shown in Fig. 554. The oscillator in this form consists of two spheres, A and B, which are placed in a glass or ebonite vessel which is filled with mineral oil, so that a spark gap of about a millimetre separates them. Two other spheres, D and E, are placed in the position shown, and are connected with the terminals of an induction coil or of a Holtz electrical machine. When the coil or machine is in action a spark passes between the knobs D and E through the spheres A and B. The passage of this spark causes electrical oscillations to be set up between A and B through the oil, which, owing to the passage of the spark, has become for the time being a conductor. The presence of the oil, since it requires a greater difference of potential to start a discharge through oil than through air, allows of the spheres A and B attaining a greater difference of potential before a spark passes than would be possible were air the dielectric between the spheres. Hence, since the quantity of energy that is stored up in the spheres before the passage of the spark is proportional to the charge, that is, to the difference of potential to which they are raised, the electrical oscillations will be more energetic, that is, of greater amplitude. In this way Righi was able to obtain electrical waves of which the wave-length was not more than 7.5 cm., while more recently, using the same disposition, but with the spheres replaced by small cylinders about 4 mm. long, Lebedew has obtained wave-lengths of less than a centimetre.

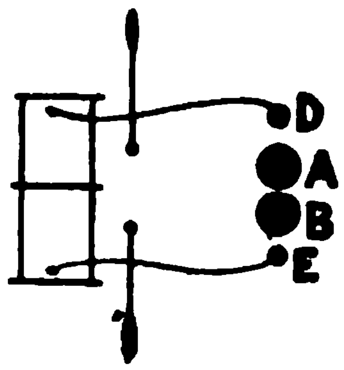


FIG. 554.

Experiment has shown that the electrical oscillations which are produced by any of the above arrangements very rapidly die out, so that the damping is very great. Thus in Fig. 555 the amplitudes of successive oscillations of an oscillator of the Hertz form is shown, the abscissæ representing time and the ordinates the difference of potential between the spheres. It will be noticed that after ten complete oscillations the amplitude is reduced to about 0.07 of its original value. Hence, since the oscillator for which this curve was drawn had a periodic time of about 3.3×10^{-8} second, the oscillations in the oscillator practically completely die out after an interval of about 10^{-6} second after they commence at the passage of the spark at the gap.

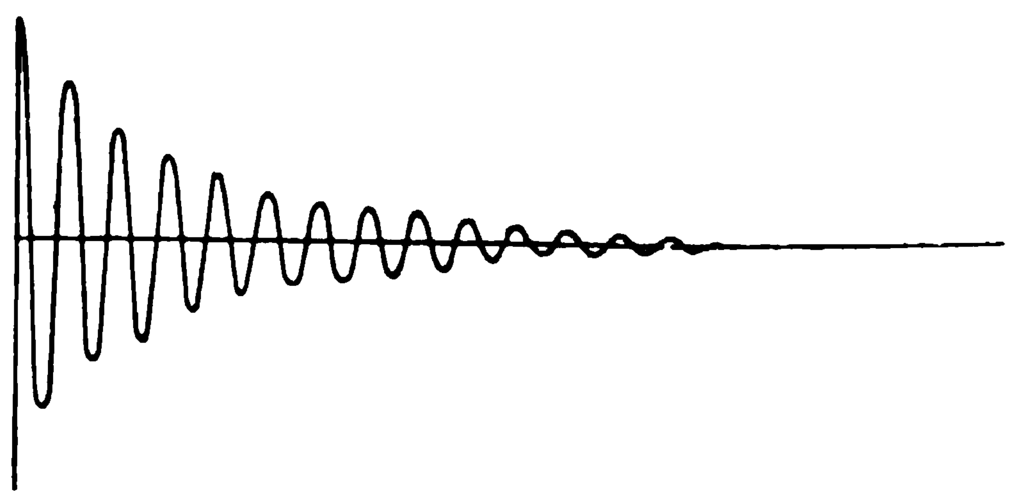


FIG. 555.

Since in this form of oscillator most of the energy is radiated in the form of waves, the power given out by an oscillator while it is acting is very considerable, amounting as it does in the case of an oscillator of the dimensions given on page 871 to about 150 horse-power during the first few vibrations. The total quantity of energy radiated is, however, small, for the time during which the oscillations are occurring is small compared to the interval between the passage of the sparks in the spark gap, so that for by far the greater proportion of time the oscillator is not acting. If we require to produce oscillations which shall last for a considerable time, it is necessary to adopt some arrangement such that while the original electrical energy stored up before the spark passes is great, the rate at which this energy is radiated during the occurrence of the oscillations shall be small. This can be accomplished by having two metal plates placed facing each other, and connected by a metal wire in which a spark gap is placed. The capacity of such an arrangement can be made large, so that the energy stored up in it is considerable, and much greater than with Hertz's oscillator, but owing to the increased capacity, the periodic time, and hence also the wave-length, is greater.

583. The Resonator.—In the preceding section we have considered the methods which have been adopted for producing electro-magnetic waves, and we must now consider some of the methods which have been employed for the detection of such waves in the space through which they may be passing. Instruments for detecting the presence of electrical waves are called resonators or receivers. The form of resonator used by Hertz is shown in Fig. 556. It consisted of a copper circle, the continuity of the copper being broken at A by a small spark gap, the length of which could be altered by means of a micrometer screw. The diameter of the circle was so chosen that the natural period of the oscillations in it was the same as the period of the oscillator.

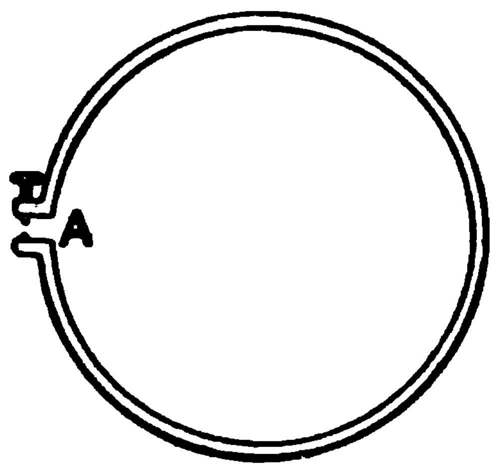


FIG. 556.

When electro-magnetic waves strike such a resonator they will induce electrical oscillations in the copper circle, and the amplitude of the oscillations set up can be measured by the length of the sparks which can be obtained at the micrometer spark gap.

Another form of resonator consists of two metal cylinders placed end to end with a spark gap between. Since the length of the sparks obtained in the resonator spark gap is very small, there is considerable difficulty even in observing them, so that to measure the maximum spark length, in order to define the amplitude of the oscillations, is almost impossible. On this account various other arrangements have been adopted for measuring the amplitude of the oscillations set up in the resonator.

Most of these methods depend on the measurement of the heat developed in a thin wire which replaces the spark gap. One such arrangement is shown in Fig. 557. The resonator consists of two metal cylinders, AB, of such a length that they are in electrical unison with the oscillator. The ends of these cylinders are connected together by a very thin wire, half being of iron and the other half of nickel. The portion *aob* is of iron, and the portion *cod* of nickel. When electrical oscillations are set up in the cylinders a current will pass backwards and forwards through the wire *aoc*, and it will thus become heated. In this way the junction at *o* between the iron and nickel wires, which is shown enlarged at D, will become heated, and a thermo-electric current will be produced in a circuit connected to the ends *b* and *d*, and this current may be measured by a galvanometer G included in the circuit.

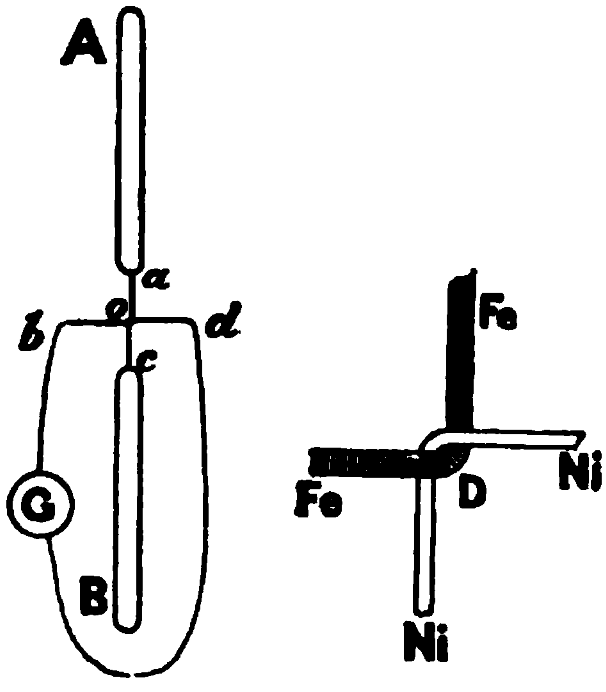


FIG. 557.

584. Stationary Electro-magnetic Waves.—Since a conductor is incapable of supporting an electro-static strain, when a Faraday tube meets a conductor, the strain which existed within the dielectric is immediately relieved, a conduction current being produced within the conductor. Owing, however, to electrical inertia the dielectric in the neighbourhood of the conducting surface overshoots its equilibrium position and so becomes the seat of an electro-static strain in the opposite sense to that in the incident tube. In this way a reflected tube will be produced at the surface of a conductor, which will move back through the dielectric. The sense of the tube will be reversed by reflection, so that here we have a case of a change of phase of half a wave-length by reflection.

Let AB, Fig. 558 (*a*), be an oscillator, and CD the section of a metal plate which is acting as a reflector of the waves sent out by the oscillator. Except in the immediate neighbourhood of the oscillator, the waves will be plane and the electrical displacement in the air will be parallel to the length of the oscillator. Let *ab*, *cd*, and *ef* represent the positions of the points where the electrical displacement is, at a given instant, a maximum, at *ab* and *ef* in one direction and at *cd* in the opposite, as indicated by the signs + and −. When the tube *ef* strikes the reflector CD a reflected tube *e′f′* will be produced in which the displacement is in the opposite direction to that in the incident tube *ef*. Hence if we indicate the reflected tubes by dotted lines we shall have near the plate a full and a dotted line in which the displacements are in opposite directions. If T is the periodic time of the oscillations, then at a time $T/4$

later the maximum displacements will have travelled through a distance equal to a quarter of the wave-length. Hence the present condition of affairs is indicated in Fig. 558 (*b*), where *e′f′* is the position of the reflected maximum of displacement corresponding to the incident one *ef*. It will now be seen that at the point L_1 the incident and reflected displacements are in the same direction, and will, therefore, produce a greater displacement than the incident wave would alone produce. A quarter of a period later the position of the maxima will be as indicated in Fig. 558 (*c*). Here at the points N and N_1 the direct and reflected displace-

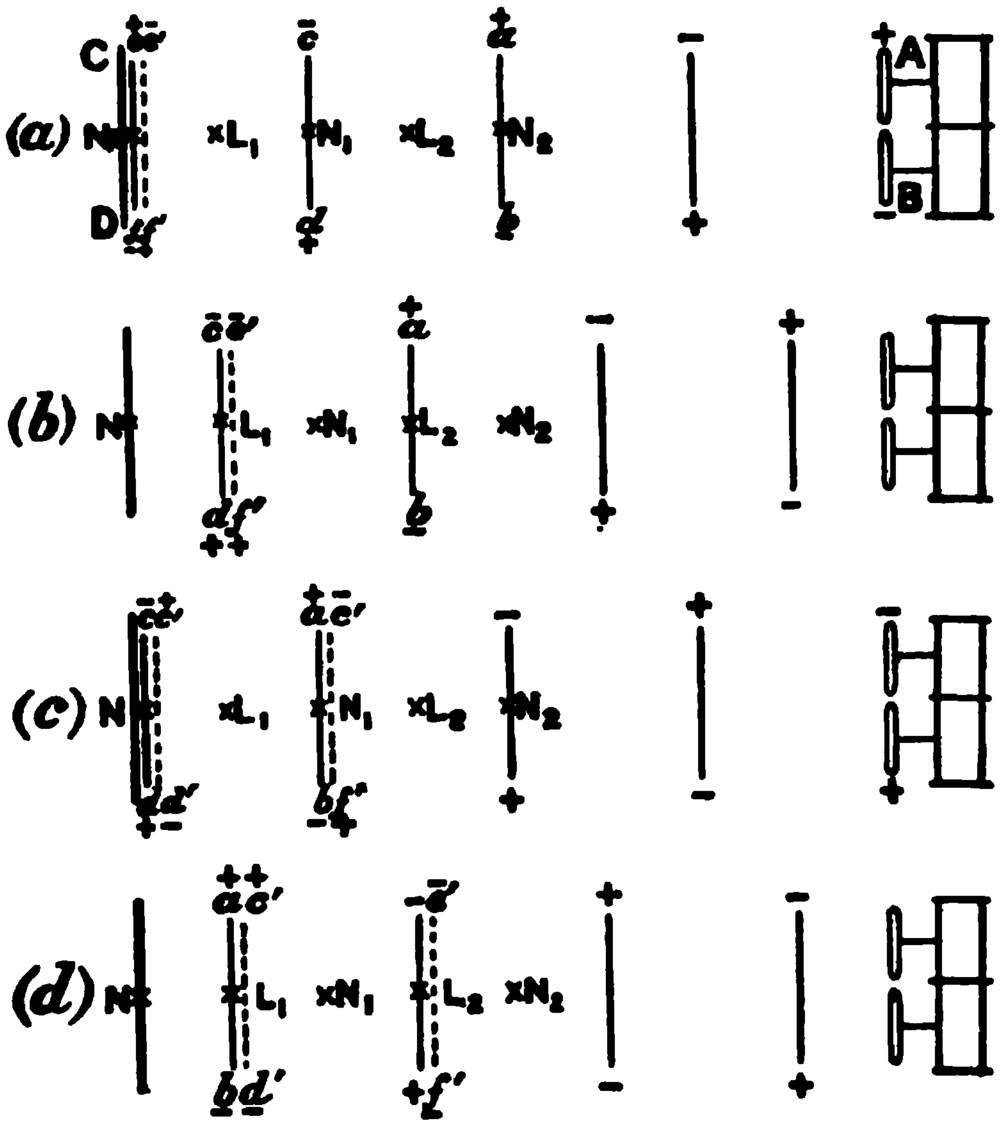


FIG. 558.

ments are in opposite directions, and hence oppose each other, so that the resultant displacement is less than would occur if the incident waves acted alone. At the end of the next quarter period the displacements due to the incident and reflected waves will be in the same direction at the points L_1 and L_2, as shown in Fig. 558 (*d*), while a quarter of a period later they will be opposed at the points N, N_1, and N_2, and so on. It will thus be seen that owing to the interference of the direct and the reflected waves a series of stationary nodes and loops will be produced, the nodes occurring at the points N, N_1, N_2, and the loops at the points L_1, L_2, for

at the points N, N_1, N_2 the incident and reflected displacements are always in opposite directions, while at the points L_1, L_2 they are always in the same direction. If then a resonator, such as that shown in Fig. 556, is moved along between the oscillator and the reflecting surface, it will respond at the points L_1 and L_2, and the galvanometer will be deflected owing to the heat developed by the oscillations in the thin connecting wire. At the points N, N_1, and N_2, however, it will not respond, since at these points the displacements due to the direct and reflected waves are in opposite directions and the galvanometer will be undeflected. Hence by moving the resonator along and noting the points where the galvanometer deflection is a maximum, the position of the loops can be found, and by noting the positions where the deflection is a minimum the nodes can in the same way be found. From these positions the wave-length of the electro-magnetic waves can at once be deduced, for it is twice the distance between two consecutive loops or nodes. Knowing the wave-length, then, if the periodic time of the oscillator is calculated from Lord Kelvin's formula, we can at once calculate the velocity with which the electro-magnetic waves travel in air. This is what Hertz was the first to do, and the value obtained from his experiments is the same as the velocity of light, if we consider the errors to which such a measurement is liable. Thus a result which Maxwell had predicted from a consideration of the manner in which one electrified body affects another through an intervening layer of dielectric, and which at the time was entirely at variance with all the accepted ideas, was, after his death, proved by Hertz to be true, and in this way has Maxwell's theory been vindicated.

It may be worth while to insist on what is actually proved by the existence of stationary waves. Their formation shows in the first place that electrical energy can be propagated through the air, a result which many other experiments also prove. Secondly, it shows that this energy takes an appreciable time to travel from the one body to the other, and that during the time between the energy leaving the oscillator and its arrival at the resonator it must exist in the intervening air.

585. The Coherer.—A very sensitive method of detecting the presence of electro-magnetic waves has been discovered by Branley. He found that a glass tube filled with loosely packed metallic filings, when included in the circuit of an electric battery and a galvanometer, was practically an insulator, so that the current could not pass, and the galvanometer was hardly, if at all, deflected. On producing electric oscillations in the neighbourhood of the circuit, however, the tube containing the filings becomes a conductor and the battery is able to drive a current through the circuit, so that the galvanometer is strongly deflected. This tube of filings forms so delicate a detecter of electrical oscillations that, as will be described later, even when the oscillator is at a distance of ten miles it will respond.

When the Branley tube of filings, or the coherer as it has been called, is employed to make measurements, the chief difficulty is to arrange the circuit so that stray electrical oscillations reflected from the walls of the room and the person of the observer do not mask the effects to be observed. For this reason it is found necessary not only to enclose the battery, galvanometer, and the connecting wires within a metallic box so as to cut off the electrical waves from these parts of the circuit, but it is necessary to pack very carefully all the joints in the box with tinfoil, or the waves will creep in in sufficient quantity to upset the indications. The only part of the circuit which is left outside the metallic box is the tube containing the filings, so that it is only when the waves fall on this tube that the circuit becomes conducting and the galvanometer is deflected. The precise way in which the electrical waves act to cause the filings to become conducting is not known, although it has been supposed that minute sparks are formed by the electrical waves between the adjacent filings, and these break down the film of condensed gas which always forms on the surface of a solid. In order to convert the tube of filings from the conducting to the non-conducting condition it is only necessary to give the tube a very slight mechanical shock, such as gently tapping with the finger.

586. Reflection, Refraction, and Polarisation of Electro-magnetic Waves.—The properties of electrical waves can be very clearly shown by means of the apparatus shown in Fig. 559. The oscillator, which is on Righi's principle, is placed along the focal line of a parabolic mirror, C, made of zinc. This mirror reflects the electrical waves in the same manner as does the reflector behind a searchlight, so that, instead of spreading out in all directions, the waves are sent in a parallel beam. The outside knobs of the oscillator are connected to the terminals of an induction coil or a Holtz electrical machine. The receiver, which consists of a coherer, is also placed along the focal line of a parabolic mirror, D, the terminals being connected with a circuit which includes a cell, E, and a galvanometer or electric bell, G, which will serve to indicate when the coherer becomes conducting. The whole of this circuit is enclosed in a metallic box to screen off stray waves. The mirror D serves to concentrate the incident waves on the receiver R, and in this way increases the sensitiveness of the apparatus.

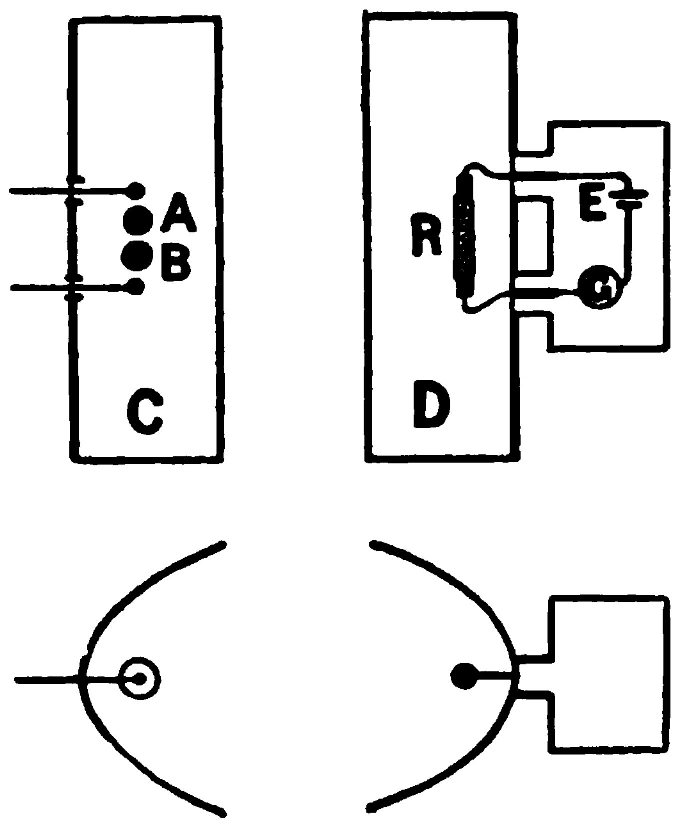


Fig. 559.

The oblique reflection of electrical waves from a metallic surface can be shown by arranging the oscillator and receiver as shown in Fig. 560, and it will be found that the receiver is only affected when the metal plate, F, is placed so that the angles of incidence and reflection are equal.

The waves produced by the oscillator are plane polarised, for the displacement is always parallel to the axis of the oscillator. On the other hand the receiver, with its parabolic reflector, only responds to waves in which the displacement is parallel to its axis. Thus the oscillator is a polariser as well as a source of the waves, while the receiver is an analyser. If, instead of using a plane sheet of metal as reflector, a grating is used made of a number of wires stretched parallel to one another on a frame, or of a number of parallel strips of tinfoil pasted on a wooden board, reflection will take place when the wires or strips are parallel to the axis of the oscillator, while in this case none of the waves will be transmitted. In fact the arrangement acts just as if it were a continuous sheet of metal. If, however, the wires are placed at right angles to the axis of the oscillator, there will be no reflection, while the waves will be transmitted just as if the wires were a dielectric. The reason for this difference is at once apparent if we remember in what way the reflected waves are produced. They are due to the induced charge on the surface of the metal caused by the incident waves. Now when the wires are parallel to the direction of the electrical displacement in the incident waves, the induced charges can be produced just as in a continuous sheet of metal, and the charges which are induced on the wires, so long as these are fairly close together, are sufficient to completely screen the portion of space behind the wires. When, however, the wires are at right angles to the direction of the displacement in the incident waves there cannot be a corresponding charge induced on the wires, for each wire being insulated from the adjacent wires, no movement of electricity can take place from one wire to the next, so that the only possible induced charge which can be produced is one on the opposide sides of each wire, and the positive induced charge on the one side of any wire will be practically neutralised by the negative charge which will be simultaneously induced on the side of the adjacent wire. A framework of conducting wires will thus act in exactly the same manner as does a plate of tourmaline in optics, reflecting waves in which the displacement is parallel to the length of the wires, but transmitting all waves in which the displacement is perpendicular to the length of the wires. The polarising effect of such a wire frame can be very clearly shown by placing the oscillator and receiver with their axes crossed as shown in Fig. 561. In this position the receiver will not be affected by

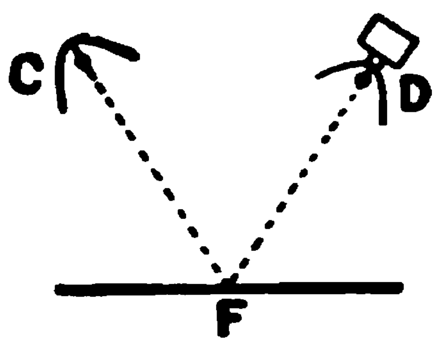


FIG. 560.

the oscillator. Neither will the receiver be influenced if the wire frame is introduced between it and the oscillator, if the length of the wires is parallel to the axis of either the oscillator or the receiver ; the reason being that when the wires are parallel to the axis of the oscillator they will not allow any of the waves to pass, and when the wires are parallel to the axis of the receiver, although the waves will now be transmitted, yet since the direction of displacement in these waves is at right angles to the axis of the receiver, they will not cause it to respond. If, however, the wires are arranged so that they are inclined at 45° to the axis of both the oscillator and the receiver, this latter will respond. The reason for this is that when the waves strike the wire frame, which is at 45° to the direction of displacement, they are resolved into two components, in which the displacements are at right angles to one another. The component in which the displacement is parallel to the wires is reflected, while that in which the displacement is at right angles to the length of the wires is transmitted. These transmitted waves, falling on the receiver, are again resolved into two components, in one of which the displacement is perpendicular to the axis of the receiver, and in the other the displacement is parallel to the axis of the receiver, and this latter will affect the receiver. The experiment corresponds to the optical experiment of introducing a doubly refracting plate between crossed Nicols. When the principal section of the crystal is parallel to the principal plane of the analyser or the polariser, no light is transmitted through the system. If, however, the principal section of the crystal is inclined at 45° to the principal planes of the Nicols, then light is transmitted through the analyser.

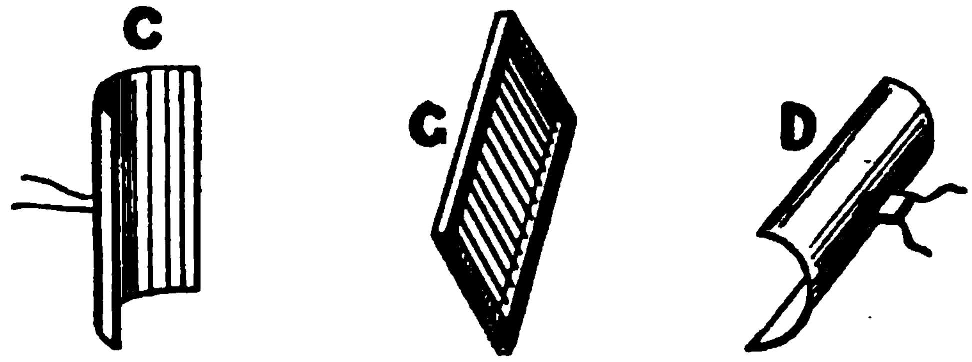


FIG. 561.

The refraction of electro-magnetic waves can be shown by means of a prism of paraffin or pitch. The prism is arranged as shown in Fig. 562, with metal screens, E and F, arranged so as to cut off any waves which do not pass through the prism. By measuring the angle of deviation through which the waves have been turned and the refracting angle of the prism, we can calculate, just as in the corresponding optical experi-

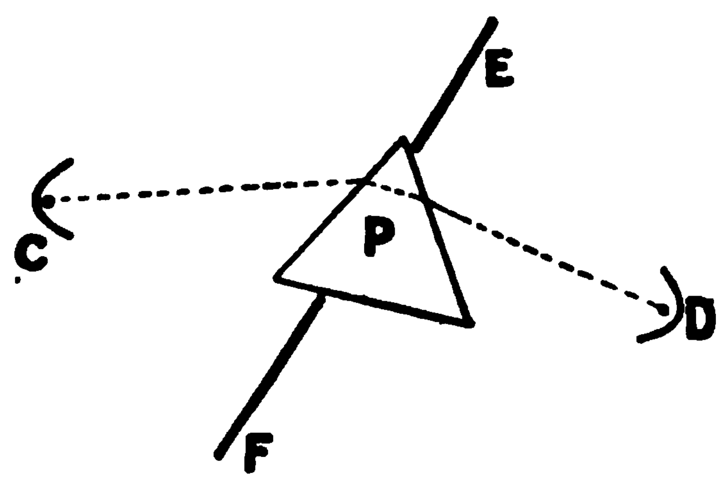


FIG. 562.

ment (§ 346), the refractive index of the material of the prism for these waves.

Experimenting in this way, the values for the refractive index of some bodies for electro-magnetic waves given in the following table have been obtained :—

REFRACTIVE INDEX FOR ELECTRO-MAGNETIC WAVES.

Substance.	Refractive Index for Wave-lengths of		
	8 mm.	6 mm.	4 mm.
Paraffin . . .	1.52	1.41	1.39
Sulphur . . .	1.80	2.01	2.00
Ebonite . . .	1.74	1.72	1.56

587. Reflection of Electro-magnetic Waves at the Surface of a Dielectric.—We have hitherto only considered the reflection of electrical waves at the surface of a conductor. When electrical waves pass from one dielectric to another, although part of the waves will be transmitted, yet a portion will be reflected. A similar phenomenon is exhibited when light passes from one transparent medium to another, and it will be well to recall briefly what peculiarities accompany such reflection. In § 407 we saw that when the incident beam was incident at a certain angle the reflected beam was plane polarised. Further, that if the incident light was plane polarised in the plane of incidence, then for a certain angle of incidence the whole of the light was reflected, while if the incident beam was polarised in a plane at right angles to the plane of incidence, none of the light was reflected. It has also been mentioned that there are two rival theories as to whether the vibrations of the ether which constitute light take place in or perpendicular to the plane of polarisation.

In the electrical case, if the axis of the oscillator, as shown in Fig. 563 (*a*), is at right angles to the plane of incidence, and if the angle of incidence, CON, at the surface of a plate of sulphur, EF, is about 60°, there will be a reflected beam, and the receiver placed at D, so that the angle DON is equal to the angle CON, will be affected. In this case the electrical displacement is perpendicular to the plane of incidence. If, however, the oscillator is arranged as in Fig. 563 (*b*), with its axis in the plane of incidence, there will be no reflected beam, and the receiver will have to be placed at D′ to be affected. Hence when the electrical displacement in the incident waves is at right angles to the plane of incidence there is reflection, but when the displacement is in the plane of incidence there is no reflected beam. Now in the case of light there

is no reflected beam when the light is polarised at right angles to the plane of incidence, that is, when, according to Fresnel's view, the displacement is in the plane of incidence. Since, then, on the electro-magnetic theory, light waves and electro-magnetic waves are the same, and only differ in wave-length, it follows that, in the case shown in Fig. 563 (*b*), the electrical waves are polarised in the plane of incidence, while in Fig. 563 (*a*) the waves are polarised perpendicular to the plane of incidence. Thus the electrical displacement takes place perpendicular to the plane of polarisation, and thus corresponds to the displacement considered by Fresnel. Since, as we have seen, the electrical displacement is always accompanied by magnetic forces which occur at right angles to the direction of the electrical displacement, it follows that the magnetic force is in the plane of polarisation. The electro-magnetic theory therefore shows that both Fresnel and Mac-

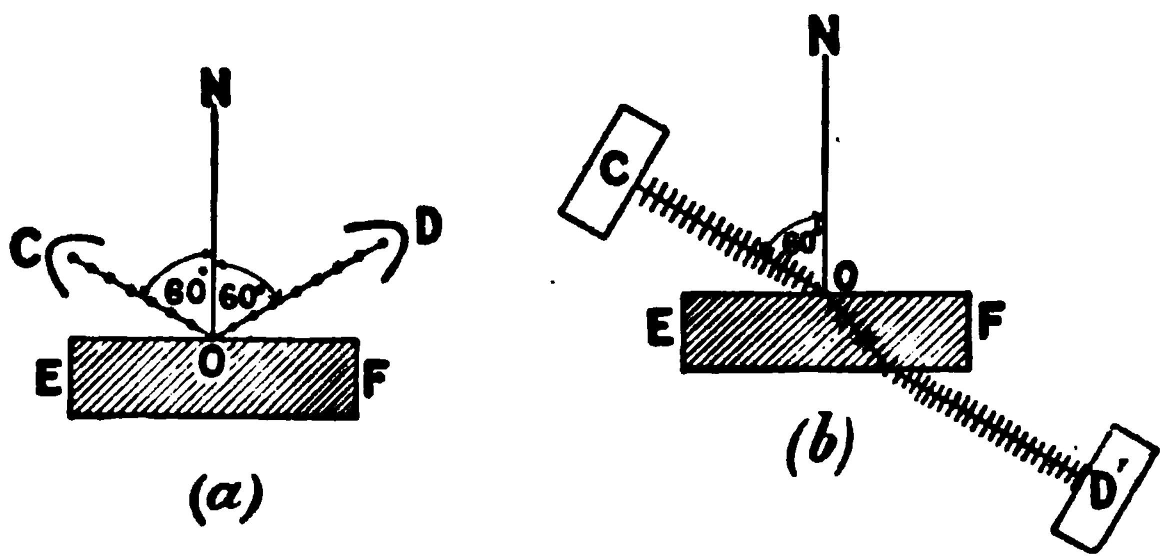


FIG. 563.

Cullagh were right as far as they went, but that neither were complete, in that in addition to the displacements they considered there is always something taking place in a perpendicular direction, both displacements however being in the wave-front, that is, at right angles to the direction in which the wave is travelling. The experiments on the reflection of electro-magnetic waves at the surface of the sulphur further show that the displacement which Fresnel considered is the electrical displacement, and that considered by MacCullagh is the magnetic.

588. Electro-magnetic Waves along Wires.—In addition to the electro-magnetic waves which are propagated in free air or other dielectric, waves can be produced in such a way that their direction of propagation is along conducting wires. The usual arrangement employed for producing these waves is shown in Fig. 564. The primary oscillator consists of two metal plates, A and B, which are connected by wires including a spark gap, O, and are also connected with the terminals of an induction coil, C. Two other metal plates, D and E, are placed opposite

the plates of the oscillator. These plates are connected with two wires, DFH and EGK, which are stretched on insulating supports at a distance from one another of about 7 cm. When electrical oscillations are produced in AB, owing to induction, oscillations will also be produced in the plates D and E. Thus waves of the same period as those in the primary oscillator will be propagated along the two wires, or rather in the dielectric between the wires. Since whenever A is positively electrified B will

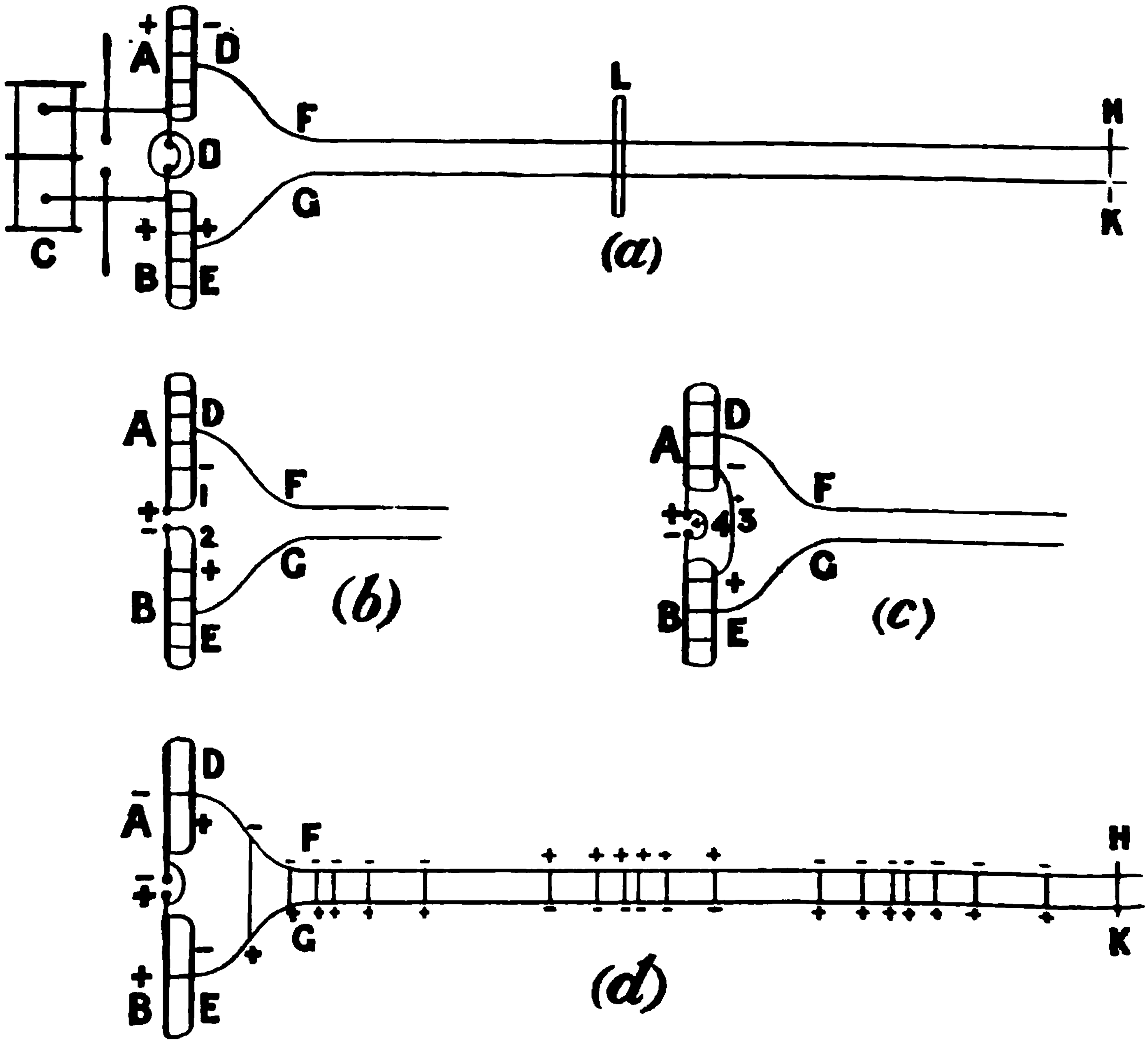


FIG. 564.

be negatively electrified, and the sign of the electrifications on D and E is always the opposite to that on the corresponding plate of the oscillator, the phase of the vibrations sent along the two wires will be opposite, that is, whenever a given point on the one wire is at its maximum positive potential, the corresponding point on the other wire will be at its maximum negative potential. If the ends of the wires are free the waves will be reflected, and if the lengths of the wires are

adjusted suitably, stationary waves will be set up owing to the interference of the direct waves with those reflected from the ends. Since it would be inconvenient to adjust the length of the wire to secure the formation of stationary waves, it is usual, instead of having the wires insulated from one another at the ends, to join them by a metallic bridge, HK, the position of which, and hence the length of the wires, can be altered.

It is instructive to consider how this arrangement works from the point of view of Faraday tubes of force. When the plates A and B are charged by the coil before the passage of a spark, the Faraday tubes will stretch somewhat as indicated in Fig. 564 (*a*). When the spark passes, due to the number of tubes stretching from one knob of the spark gap to the other being so great that the electrical stress, of which the tubes are simply the graphic representation, overcomes the dielectric strength of the air, the tubes which stretch across in the path of the spark, since along this path the air becomes a conductor, will be able to shrink to nothing. The disappearance of these tubes will allow some of the tubes which were crowded out into the intervening space, owing to the transverse repulsion between the tubes, to move towards the gap as shown at (*b*). The two tubes marked 1 and 2 will, since they are turned in opposite senses, attract one another, and they will first fuse together and then separate into two tubes, 3 and 4, as shown at (*c*). Tube 4 will rush into the conducting spark gap to keep up the discharge, while tube 3 will move through the dielectric separating the wires, with its positive end on the wire EGK, and its negative end on the wire DFH. As these tubes move away, another pair of tubes will move out from the space between the plates, and will go through the same operation. The process will not, however, stop when the plates have lost all their tubes, that is, when all the Faraday tubes have moved out from between the plates, but, owing to inertia, a dielectric displacement will be produced in the opposite direction, that is, tubes will appear which stretch in the opposite sense to the original tubes. The crowding in of these tubes between the plates may be regarded as simply the passing out from between the plates of more tubes in the original sense than there were originally there. Then these tubes will go through the same series of operations as the others did, and tubes will be propagated along the wires which have their positive ends on the wire DFH, and their negative ends on EGK. Thus there will be a number of sets of tubes, such as are shown at (*d*), travelling along the wires. When a tube reaches the bridge HK, owing to inertia it will not simply shrink to nothing, but will stretch out again and travel back; but if, when it approached the bridge, its positive end was on the wire FH, when travelling back its negative end will be on this wire. We shall thus have a train of reflected tubes, and these, together with the direct tubes which are travelling towards the bridge, will produce a system of nodes and loops, for at certain points on the wires the tubes which

reach there on their way to the bridge will always be of the opposite sign to those which reach this point after reflection, and hence the two sets of tubes will neutralise each other. At other points the two sets of tubes will always be in the same sense, and will therefore produce a loop. The experiment is exactly analogous to Kundt's method of determining the velocity of sound in gases, as described in § 314; the oscillator here corresponding to the vibrating rod which in the acoustical experiment produces the vibrations of the gas in the tube. The bridge here corresponds to the closed end of the tube, for in both cases they correspond to a node, that is, in the one case to a point where the movement of the air is a minimum, and in the other to a point where the electrical displacement is a minimum. The electrical displacement at the bridge is a minimum, for there the potential of the two wires is always the same, so that there cannot be any electro-static force, and hence no displacement. Also, just as in Kundt's experiment, the stationary waves set up in the tube are very much more intense if the length is adjusted so that it is a multiple of the half wave-length of the note given by the rod in the gas, so in the electrical case the amplitude of the stationary waves set up in the wires is much increased if the position of the bridge is altered till the length of each wire is some multiple of the half wave-length of the electrical waves in the wires produced by the oscillator.

The position of the loops on the wires can be determined by placing a Geissler tube, L (§ 381), across the wires as a bridge. If the tube is moved along it will glow brightly at the loops, but will be dark at the nodes. In this way the wave-length of the oscillations in the wires can be measured. Then, if the periodic time of the primary oscillations is known, the velocity with which the oscillations travel, when conducted in this way along wires, can be calculated. By this method, as well as by a direct comparison between the velocity in free air with that along a wire, it has been proved that the velocity is the same when the waves are propagated in a free dielectric as when they are conducted along a wire. Also, it is found that the velocity, while it is independent of the material of the wires, depends on the specific inductive capacity of the dielectric which surrounds the wires. This result is a conclusive proof that the energy travels, not along the wires, but through the dielectric which surrounds them, as is indicated by Maxwell's theory.

589. Telegraphy without Connecting Wires.—Within the last year or two much attention has been devoted to the employment of electro-magnetic waves as a means of transmitting signals from one place to another without the necessity for a metallic wire connecting the two. There are a number of arrangements which have been tried, and the whole subject is still (1899) in a very experimental stage. Some of the most successful attempts have been made by Marconi, who has transmitted signals from Alum Bay to Bournemouth, a distance of eighteen miles. As transmitter he uses an oscillator of Righi's form, giving waves of

about 120 cm. wave-length. As a receiver he uses a coherer, which is placed in series with a relay; this relay working a sounder on which the Morse signals are received. In order to make the coherer lose its conductivity when the waves stop, an electro-magnet works a small hammer which is continually tapping the tube. In the case of transmission over distances greater than a mile or two, a collector, consisting of a vertical wire attached to a pole, is used to conduct the waves to the coherer.

INDEX

Morse alphabet, instructions 775

THE END

Printed by Ballantyne, Hanson & Co.
Edinburgh & London

A

CLASSIFIED CATALOGUE

OF

SCIENTIFIC WORKS

PUBLISHED BY

MESSRS. LONGMANS, GREEN, & CO.

LONDON: 39 PATERNOSTER ROW, E.C.

NEW YORK: 91 & 93 FIFTH AVENUE.

BOMBAY: 32 HORNBY ROAD.

CONTENTS.

CHEMISTRY.

CROOKES.—SELECT METHODS IN CHEMICAL ANALYSIS, chiefly Inorganic. By Sir WILLIAM CROOKES, F.R.S., etc. Third Edition, Rewritten and Enlarged. With 67 Woodcuts. 8vo., 21*s*. net.

FURNEAUX.—ELEMENTARY CHEMISTRY, Inorganic and Organic. By W. FURNEAUX, F.R.C.S., Lecturer on Chemistry, London School Board. With 65 Illustrations and 155 Experiments. Crown 8vo., 2*s*. 6*d*.

GARRETT AND HARDEN.—AN ELEMENTARY COURSE OF PRACTICAL ORGANIC CHEMISTRY. By F. C. GARRETT, M.Sc. (Vict. et Dunelm.), Assistant Lecturer and Demonstrator in Chemistry, the Durham College of Science, Newcastle-on-Tyne; and ARTHUR HARDEN, M.Sc. (Vict.), Ph.D., Assistant Lecturer and Demonstrator in Chemistry, the Owens College, Manchester. With 14 Illustrations. Crown 8vo., 2*s*.

JAGO.—Works by W. JAGO, F.C.S., F.I.C.

INORGANIC CHEMISTRY, THEORETICAL AND PRACTICAL. With an Introduction to the Principles of Chemical Analysis. Inorganic and Organic. With 63 Woodcuts and numerous Questions and Exercises. Fcp. 8vo., 2*s*. 6*d*.

AN INTRODUCTION TO PRACTICAL INORGANIC CHEMISTRY. Crown 8vo., 1*s*. 6*d*.

INORGANIC CHEMISTRY, THEORETICAL AND PRACTICAL. A Manual for Students in Advanced Classes of the Science and Art Department. With Plate of Spectra and 78 Woodcuts. Crown 8vo., 4*s*. 6*d*.

KOLBE.—A SHORT TEXT-BOOK OF INORGANIC CHEMISTRY. By Dr. HERMANN KOLBE. Translated and Edited by T. S. HUMPIDGE, Ph.D. With 66 Illustrations. Crown 8vo., 8*s*. 6*d*.

MENDELÉEFF.—THE PRINCIPLES OF CHEMISTRY. By D. MENDELÉEFF. Translated from the Russian (Sixth Edition) by GEORGE KAMENSKY, A.R.S.M., of the Imperial Mint, St. Petersburg; and Edited by T. A. LAWSON, B.Sc., Ph.D., Fellow of the Institute of Chemistry. With 96 Diagrams and Illustrations. 2 vols. 8vo., 36*s*.

MEYER.—OUTLINES OF THEORETICAL CHEMISTRY. By LOTHAR MEYER, Professor of Chemistry in the University of Tübingen. Translated by Professors P. PHILLIPS BEDSON, D.Sc., and W. CARLETON WILLIAMS, B.Sc. 8vo., 9*s*.

MILLER.—INTRODUCTION TO THE STUDY OF INORGANIC CHEMISTRY. By W. ALLEN MILLER, M.D., LL.D. With 71 Woodcuts. Fcp. 8vo., 3*s*. 6*d*.

CHEMISTRY—*Continued.*

MUIR.—A COURSE OF PRACTICAL CHEMISTRY. By M. M. P. MUIR, M.A., Fellow and Prælector in Chemistry of Gonville and Caius College, Cambridge. (3 Parts.)

Part I. Elementary. Crown 8vo., 4*s.* 6*d.*

Part II. Intermediate. Crown 8vo., 4*s.* 6*d.*

Part III. [*In preparation.*

NEWTH.—Works by G. S. NEWTH, F.I.C., F.C.S., Demonstrator in the Royal College of Science, London; Assistant Examiner in Chemistry, Science and Art Department.

CHEMICAL LECTURE EXPERIMENTS. With 224 Diagrams. Crown 8vo., 10*s.* 6*d.*

CHEMICAL ANALYSIS, QUANTITATIVE AND QUALITATIVE. With 100 Illustrations. Crown 8vo., 6*s.* 6*d.*

A TEXT-BOOK OF INORGANIC CHEMISTRY. With 146 Illustrations. Crown 8vo., 6*s.* 6*d.*

ELEMENTARY PRACTICAL CHEMISTRY. With 108 Illustrations and 254 Experiments. Crown 8vo., 2*s.* 6*d.*

OSTWALD.—SOLUTIONS. By W. OSTWALD, Professor of Chemistry in the University of Leipzig. Being the Fourth Book, with some additions, of the Second Edition of Oswald's 'Lehrbuch der allgemeinen Chemie'. Translated by M. M. PATTISON MUIR, Fellow and Prælector in Chemistry of Gonville and Caius College, Cambridge. 8vo., 10*s.* 6*d.*

REYNOLDS.—EXPERIMENTAL CHEMISTRY FOR JUNIOR STUDENTS. By J. EMERSON REYNOLDS, M.D., F.R.S., Professor of Chemistry, University of Dublin. Fcp. 8vo., with numerous Woodcuts.

Part I. Introductory. Fcp. 8vo., 1*s.* 6*d.*

Part II. Non-Metals, with an Appendix on Systematic Testing for Acids. Fcp. 8vo., 2*s.* 6*d.*

Part III. Metals, and Allied Bodies. Fcp. 8vo., 3*s.* 6*d.*

Part IV. Carbon Compounds. Fcp. 8vo., 4*s.*

SHENSTONE.—Works by W. A. SHENSTONE, Lecturer on Chemistry in Clifton College.

THE METHODS OF GLASS-BLOWING. For the use of Physical and Chemical Students. With 42 Illustrations. Crown 8vo., 1*s.* 6*d.*

A PRACTICAL INTRODUCTION TO CHEMISTRY. Intended to give a Practical acquaintance with the Elementary Facts and Principles of Chemistry. With 25 Illustrations. Crown 8vo., 2*s.*

CHEMISTRY—*Continued.*

THORNTON AND PEARSON.—NOTES ON VOLUMETRIC ANALYSIS. By ARTHUR THORNTON, M.A., Senior Science Master, Bradford Grammar School; and MARCHANT PEARSON, B.A., Assistant Science Master, Bradford Grammar School. Medium 8vo., 2*s.*

THORPE.—Works by T. E. THORPE, B.Sc. (Vict.), Ph.D., F.R.S., Professor of Chemistry in the Royal College of Science, South Kensington. Assisted by Eminent Contributors.

A DICTIONARY OF APPLIED CHEMISTRY. 3 vols. 8vo. Vols. I. and II., 42*s.* each. Vol. III., 63*s.*

QUANTITATIVE CHEMICAL ANALYSIS. With 88 Woodcuts. Fcp. 8vo., 4*s.* 6*d.*

THORPE AND MUIR.—QUALITATIVE CHEMICAL ANALYSIS AND LABORATORY PRACTICE. By T. E. THORPE, Ph.D., D.Sc., F.R.S., and M. M. PATTISON MUIR, M.A. With Plate of Spectra and 57 Woodcuts. Fcp. 8vo., 3*s.* 6*d.*

TILDEN.—Works by WILLIAM A. TILDEN, D.Sc. London, F.R.S., Professor of Chemistry in the Royal College of Science, South Kensington.

A SHORT HISTORY OF THE PROGRESS OF SCIENTIFIC CHEMISTRY IN OUR OWN TIMES. Crown 8vo., 5*s.* net.

INTRODUCTION TO THE STUDY OF CHEMICAL PHILOSOPHY. The Principles of Theoretical and Systematic Chemistry. With 5 Woodcuts. With or without the ANSWERS of Problems. Fcp. 8vo., 4*s.* 6*d.*

PRACTICAL CHEMISTRY. The principles of Qualitative Analysis. Fcp. 8vo., 1*s.* 6*d.*

HINTS ON THE TEACHING OF ELEMENTARY CHEMISTRY IN SCHOOLS AND SCIENCE CLASSES. With 7 Illustrations. Crown 8vo., 2*s.*

WHITELEY.—Works by R. LLOYD WHITELEY, F.I.C., Principal of the Municipal Science School, West Bromwich.

CHEMICAL CALCULATIONS. With Explanatory Notes, Problems and Answers, specially adapted for use in Colleges and Science Schools. With a Preface by Professor F. CLOWES, D.Sc. (Lond.), F.I.C. Crown 8vo., 2*s.*

ORGANIC CHEMISTRY: the Fatty Compounds. With 45 Illustrations. Crown 8vo., 3*s.* 6*d.*

PHYSICS, ETC.

EARL.—THE ELEMENTS OF LABORATORY WORK: a Course of Natural Science. By A. G. EARL, M.A., F.C.S., late Scholar of Christ's College, Cambridge. With 57 Diagrams and numerous Exercises and Questions. Crown 8vo., 4*s.* 6*d.*

GANOT.—Works by PROFESSOR GANOT. Translated and Edited by E. ATKINSON, Ph.D., F.C.S.

ELEMENTARY TREATISE ON PHYSICS, Experimental and Applied. With 9 Coloured Plates and Maps, and 1057 Woodcuts, and Appendix of Problems and Examples with Answers. Crown 8vo., 15*s.*

NATURAL PHILOSOPHY FOR GENERAL READERS AND YOUNG PERSONS. With 7 Plates, 624 Woodcuts, and an Appendix of Questions. Crown 8vo., 7*s.* 6*d.*

GLAZEBROOK AND SHAW.—PRACTICAL PHYSICS. By R. T. GLAZEBROOK, M.A., F.R.S., and W. N. SHAW, M.A. With 134 Woodcuts. Fcp. 8vo., 7*s.* 6*d.*

GUTHRIE.—MOLECULAR PHYSICS AND SOUND. By F. GUTHRIE, Ph.D. With 91 Diagrams. Fcp. 8vo., 1*s.* 6*d.*

HELMHOLTZ.—POPULAR LECTURES ON SCIENTIFIC SUBJECTS. By HERMANN VON HELMHOLTZ. Translated by E. ATKINSON, Ph.D., F.C.S., formerly Professor of Experimental Science, Staff College. With 68 Illustrations. 2 vols., crown 8vo., 3*s.* 6*d.* each.

CONTENTS.—Vol. I.—The Relation of Natural Science to Science in General—Goethe's Scientific Researches—The Physiological Causes of Harmony in Music—Ice and Glaciers—The Interaction of the Natural Forces—The Recent Progress of the Theory of Vision—The Conservation of Force—The Aim and Progress of Physical Science.

CONTENTS.—Vol. II.—Gustav Magnus. In Memoriam—The Origin and Significance of Geometrical Axioms—The Relation of Optics to Painting—The Origin of the Planetary System—Thought in Medicine—Academic Freedom in German Universities—Hermann Von Helmholtz—An Autobiographical Sketch.

HENDERSON.—ELEMENTARY PHYSICS. By JOHN HENDERSON, D.Sc. (Edin.), A.I.E.E., Physics Department, Borough Road Polytechnic. Crown 8vo., 2*s.* 6*d.*

MEYER.—THE KINETIC THEORY OF GASES. Elementary Treatise, with Mathematical Appendices. By Dr. OSKAR EMIL MEYER, Professor of Physics at the University of Breslau. Second Revised Edition. Translated by ROBERT E. BAYNES, M.A., Student of Christ Church, Oxford, and Dr. Lee's Reader in Physics. 8vo.

VAN 'T HOFF.—THE ARRANGEMENT OF ATOMS IN SPACE. By J. H. VAN T'HOFF. Second, Revised, and Enlarged Edition. With a Preface by JOHANNES WISLICENUS, Professor of Chemistry at the University of Leipzig; and an Appendix 'Stereo-chemistry among Inorganic Substances,' by ALFRED WERNER, Professor of Chemistry at the University of Zürich. Translated and Edited by ARNOLD EILOART. Crown 8vo., 6*s.* 6*d.*

PHYSICS, ETC.—*Continued.*

WATSON.—ELEMENTARY PRACTICAL PHYSICS: a Laboratory Manual for Use in Organised Science Schools. By W. WATSON, B.Sc. Assistant Professor in Physics in the Royal College of Science, London; Assistant Examiner in Physics, Science and Art Department. With 120 Illustrations and 193 Exercises. Crown 8vo., 2*s.* 6*d.*

WORTHINGTON.—A FIRST COURSE OF PHYSICAL LABORATORY PRACTICE. Containing 264 Experiments. By A. M. WORTHINGTON, M.A., F.R.S. With Illustrations. Crown 8vo., 4*s.* 6*d.*

WRIGHT.—ELEMENTARY PHYSICS. By MARK R. WRIGHT, M.A., Professor of Normal Education, Durham College of Science. With 242 Illustrations. Crown 8vo., 2*s.* 6*d.*

MECHANICS, DYNAMICS, STATICS, HYDRO-STATICS, ETC.

BALL.—A CLASS-BOOK OF MECHANICS. By Sir R. S. BALL, LL.D. 89 Diagrams. Fcp. 8vo., 1*s.* 6*d.*

GELDARD.—STATICS AND DYNAMICS. By C. GELDARD, M.A., formerly Scholar of Trinity College, Cambridge. Crown 8vo., 5*s.*

GOODEVE.—Works by T. M. GOODEVE, M.A., formerly Professor of Mechanics at the Normal School of Science, and the Royal School of Mines.

THE ELEMENTS OF MECHANISM. With 357 Woodcuts. Crown 8vo., 6*s.*

PRINCIPLES OF MECHANICS. With 253 Woodcuts and numerous Examples. Crown 8vo., 6*s.*

A MANUAL OF MECHANICS: an Elementary Text-Book for Students of Applied Mechanics. With 138 Illustrations and Diagrams, and 188 Examples taken from the Science Department Examination Papers, with Answers. Fcp. 8vo., 2*s.* 6*d.*

GOODMAN.—MECHANICS APPLIED TO ENGINEERING. By JOHN GOODMAN, Wh.Sch., A.M.I.C.E., M.I.M.E., Professor of Engineering in the Yorkshire College, Leeds (Victoria University). With 620 Illustrations and numerous examples. Crown 8vo., 7*s.* 6*d.* net.

GRIEVE.—LESSONS IN ELEMENTARY MECHANICS. By W. H. GRIEVE, late Engineer, R.N., Science Demonstrator for the London School Board, etc.

Stage 1. With 165 Illustrations and a large number of Examples. Fcp. 8vo., 1*s.* 6*d.*
Stage 2. With 122 Illustrations. Fcp. 8vo., 1*s.* 6*d.*
Stage 3. With 103 Illustrations. Fcp. 8vo., 1*s.*

MECHANICS, DYNAMICS, STATICS, HYDROSTATICS, ETC.—*Continued.*

MAGNUS.—Works by SIR PHILIP MAGNUS, B.Sc., B.A.

LESSONS IN ELEMENTARY MECHANICS. Introductory to the study of Physical Science. Designed for the Use of Schools, and of Candidates for the London Matriculation and other Examinations. With numerous Exercises, Examples, Examination Questions, and Solutions, etc., from 1870-1895. With Answers, and 131 Woodcuts. Fcp. 8vo., 3*s.* 6*d.*

Key for the use of Teachers only, price 5*s.* 3½*d.*

HYDROSTATICS AND PNEUMATICS. Fcp. 8vo., 1*s.* 6*d.*; or, with Answers, 2*s.* The Worked Solutions of the Problems, 2*s.*

ROBINSON.—Works by the Rev. J. L. ROBINSON, M.A.

ELEMENTS OF DYNAMICS (Kinetics and Statics). With numerous Exercises. A Text-book for Junior Students. Crown 8vo., 6*s.*

A FIRST BOOK IN STATICS AND DYNAMICS. With numerous Examples and Answers. Crown 8vo, 3*s.* 6*d.*

Sold separately: Statics, 2*s.*; Dynamics, 2*s.*

SMITH.—Works by J. HAMBLIN SMITH, M.A.

ELEMENTARY STATICS. Crown 8vo., 3*s.*

ELEMENTARY HYDROSTATICS. Crown 8vo., 3*s.*

KEY TO STATICS AND HYDROSTATICS. Crown 8vo., 6*s.*

TARLETON.—AN INTRODUCTION TO THE MATHEMATICAL THEORY OF ATTRACTION. By FRANCIS A. TARLETON, LL.D., Sc.D., Fellow of Trinity College, and Professor of Natural Philosophy in the University of Dublin. Crown 8vo., 10*s.* 6*d.*

TAYLOR.—Works by J. E. TAYLOR, M.A., B.Sc. (Lond.), Head Master of the Central Higher Grade and Science School, Sheffield.

THEORETICAL MECHANICS, including Hydrostatics and Pneumatics. With 175 Diagrams and Illustrations, and 522 Examination Questions and Answers. Crown 8vo., 2*s.* 6*d.*

THEORETICAL MECHANICS—SOLIDS. With 163 Illustrations, 120 Worked Examples and over 500 Examples from Examination Papers, etc. Crown 8vo., 2*s.* 6*d.*

THEORETICAL MECHANICS.—FLUIDS. With 122 Illustrations, numerous Worked Examples, and about 500 Examples from Examination Papers, etc. Crown 8vo., 2*s.* 6*d.*

THORNTON.—THEORETICAL MECHANICS—SOLIDS. Including Kinematics, Statics and Kinetics. By ARTHUR THORNTON, M.A., F.R.A.S. With 200 Illustrations, 130 Worked Examples, and over 900 Examples from Examination Papers, etc. Crown 8vo., 4*s.* 6*d.*

MECHANICS, DYNAMICS, STATICS, HYDROSTATICS, ETC.—*Continued.*

TWISDEN.—Works by the Rev. JOHN F. TWISDEN, M.A.

PRACTICAL MECHANICS; an Elementary Introduction to their Study. With 855 Exercises, and 184 Figures and Diagrams. Crown 8vo., 10*s.* 6*d.*

THEORETICAL MECHANICS. With 172 Examples, numerous Exercises, and 154 Diagrams. Crown 8vo., 8*s.* 6*d.*

WILLIAMSON.—INTRODUCTION TO THE MATHEMATICAL THEORY OF THE STRESS AND STRAIN OF ELASTIC SOLIDS. By BENJAMIN WILLIAMSON, D.Sc., F.R.S. Crown 8vo., 5*s.*

WILLIAMSON AND TARLETON.—AN ELEMENTARY TREATISE ON DYNAMICS. Containing Applications to Thermodynamics, with numerous Examples. By BENJAMIN WILLIAMSON, D.Sc., F.R.S., and FRANCIS A. TARLETON, LL.D. Crown 8vo., 10*s.* 6*d.*

WORTHINGTON.—DYNAMICS OF ROTATION: an Elementary Introduction to Rigid Dynamics. By A. M. WORTHINGTON, M.A., F.R.S. Crown 8vo., 4*s.* 6*d.*

OPTICS AND PHOTOGRAPHY.

ABNEY.—A TREATISE ON PHOTOGRAPHY. By Captain W. DE WIVELESLIE ABNEY, F.R.S., Director for Science in the Science and Art Department. With 115 Woodcuts. Fcp. 8vo., 3*s.* 6*d.*

GLAZEBROOK.—PHYSICAL OPTICS. By R. T. GLAZEBROOK, M.A., F.R.S., Principal of University College, Liverpool. With 183 Woodcuts of Apparatus, etc. Fcp. 8vo., 6*s.*

WRIGHT.—OPTICAL PROJECTION: a Treatise on the Use of the Lantern in Exhibition and Scientific Demonstration. By LEWIS WRIGHT, Author of 'Light: a Course of Experimental Optics'. With 232 Illustrations. Crown 8vo., 6*s.*

SOUND, LIGHT, HEAT, AND THERMODYNAMICS.

CUMMING.—HEAT TREATED EXPERIMENTALLY. By LINNÆUS CUMMING, M.A. With 192 Illustrations. Crown 8vo., 4*s.* 6*d.*

DAY.—NUMERICAL EXAMPLES IN HEAT. By R. E. DAY, M.A. Fcp. 8vo., 3*s.* 6*d.*

SOUND, LIGHT, HEAT, AND THERMODYNAMICS—*Continued.*

EMTAGE.—LIGHT. By W. T. A. EMTAGE, M.A. With 232 Illustrations. Crown 8vo., 6*s.*

HELMHOLTZ.—ON THE SENSATIONS OF TONE AS A PHYSIOLOGICAL BASIS FOR THE THEORY OF MUSIC. By HERMANN VON HELMHOLTZ. Royal 8vo., 28*s.*

MADAN.—AN ELEMENTARY TEXT-BOOK ON HEAT. For the Use of Schools. By H. G. MADAN, M.A., F.C.S., Fellow of Queen's College, Oxford; late Assistant Master at Eton College. Crown 8vo., 9*s.*

MAXWELL.—THEORY OF HEAT. By J. CLERK MAXWELL, M.A., F.R.SS., L. and E. With Corrections and Additions by Lord RAYLEIGH. With 38 Illustrations. Fcp. 8vo., 4*s.* 6*d.*

SMITH.—THE STUDY OF HEAT. By J. HAMBLIN SMITH, M.A., of Gonville and Caius College, Cambridge. Crown 8vo., 3*s.*

TYNDALL.—Works by JOHN TYNDALL, D.C.L., F.R.S. See p. 28.

WORMELL.—A CLASS-BOOK OF THERMODYNAMICS. By RICHARD WORMELL, B.Sc., M.A. Fcp. 8vo., 1*s.* 6*d.*

WRIGHT.—Works by MARK R. WRIGHT, M.A.

SOUND, LIGHT, AND HEAT. With 160 Diagrams and Illustrations. Crown 8vo., 2*s.* 6*d.*

ADVANCED HEAT. With 136 Diagrams and numerous Examples and Examination Papers. Crown 8vo., 4*s.* 6*d.*

STEAM, OIL, AND GAS ENGINES.

BALE.—A HAND-BOOK FOR STEAM USERS; being Rules for Engine Drivers and Boiler Attendants, with Notes on Steam Engine and Boiler Management and Steam Boiler Explosions. By M. POWIS BALE, M.I.M.E., A.M.I.C.E. Fcp. 8vo., 2*s.* 6*d.*

CLERK.—THE GAS AND OIL ENGINE. By DUGALD CLERK, Associate Member of the Institution of Civil Engineers, Fellow of the Chemical Society, Member of the Royal Institution, Fellow of the Institute of Patent Agents. With 228 Illustrations. 8vo., 15*s.*

HOLMES.—THE STEAM ENGINE. By GEORGE C. V. HOLMES, Whitworth Scholar, Secretary of the Institution of Naval Architects. With 212 Woodcuts. Fcp. 8vo., 6*s.*

STEAM, OIL, AND GAS ENGINES—*Continued.*

NORRIS.—A PRACTICAL TREATISE ON THE 'OTTO' CYCLE GAS ENGINE. By WILLIAM NORRIS, M.I.Mech.E. With 207 Illustrations. 8vo., 10*s.* 6*d.*

RIPPER.—Works by WILLIAM RIPPER, Professor of Mechanical Engineering in the Sheffield Technical School.

STEAM. With 142 Illustrations. Crown 8vo., 2*s.* 6*d.*

STEAM ENGINE THEORY AND PRACTICE. 8vo.

SENNETT AND ORAM.—THE MARINE STEAM ENGINE: A Treatise for Engineering Students, Young Engineers and Officers of the Royal Navy and Mercantile Marine. By the late RICHARD SENNETT, Engineer-in-Chief of the Navy, etc.; and HENRY J. ORAM, Senior Engineer Inspector at the Admiralty, Inspector of Machinery in H.M. Fleet, etc. With 412 Diagrams. 8vo., 21*s.*

STROMEYER.—MARINE BOILER MANAGEMENT AND CONSTRUCTION. Being a Treatise on Boiler Troubles and Repairs, Corrosion, Fuels, and Heat, on the properties of Iron and Steel, on Boiler Mechanics, Workshop Practices, and Boiler Design. By C. E. STROMEYER, Member of the Institute of Naval Architects, etc. 8vo., 18*s.* net.

BUILDING CONSTRUCTION.

ADVANCED BUILDING CONSTRUCTION. By the Author of 'Rivingtons' Notes on Building Construction'. With 385 Illustrations. Crown 8vo., 4*s.* 6*d.*

BURRELL.—BUILDING CONSTRUCTION. By EDWARD J. BURRELL, Second Master of the People's Palace Technical School, London. With 303 Working Drawings. Crown 8vo., 2*s.* 6*d.*

SEDDON.—BUILDER'S WORK AND THE BUILDING TRADES. By Col. H. C. SEDDON, R.E., late Superintending Engineer, H.M.'s Dockyard, Portsmouth; Examiner in Building Construction, Science and Art Department, South Kensington. With numerous Illustrations. Medium 8vo., 16*s.*

RIVINGTONS' COURSE OF BUILDING CONSTRUCTION.

NOTES ON BUILDING CONSTRUCTION. Arranged to meet the requirements of the syllabus of the Science and Art Department of the Committee of Council on Education, South Kensington. Medium 8vo.

Part I. Elementary Stage. With 552 Woodcuts, 10*s.* 6*d.*

Part II. Advanced Stage. With 479 Woodcuts, 10*s.* 6*d.*

Part III. Materials. Course for Honours. With 188 Woodcuts, 21*s.*

Part IV. Calculations for Building Structures. Course for Honours. With 597 Woodcuts, 15*s.*

ELECTRICITY AND MAGNETISM.

CARUS-WILSON.—ELECTRO-DYNAMICS: the Direct-Current Motor. By CHARLES ASHLEY CARUS-WILSON, M.A. Cantab., Professor of Electrical Engineering at McGill University, Montreal. With 71 Diagrams, and a Series of Problems, with Answers. Crown 8vo., 7*s.* 6*d.*

CUMMING.—ELECTRICITY TREATED EXPERIMENTALLY. By LINNÆUS CUMMING, M.A. With 242 Illustrations. Cr. 8vo., 4*s.* 6*d.*

DAY.—EXERCISES IN ELECTRICAL AND MAGNETIC MEASUREMENTS, with Answers. By R. E. DAY. 12mo., 3*s.* 6*d.*

DU BOIS.—THE MAGNETIC CIRCUIT IN THEORY AND PRACTICE. By Dr. H. DU BOIS, Privatdocent in the University of Berlin. Translated by E. ATKINSON, Ph.D. With 94 Illustrations. 8vo., 12*s.* net.

EBERT.—MAGNETIC FIELDS OF FORCE. By H. EBERT, Professor of Physics in the University of Kiel. With 93 Illustrations. 8vo., 10*s.* 6*d.* net.

GORE.—THE ART OF ELECTRO-METALLURGY, including all known Processes of Electro-Deposition. By G. GORE, LL.D., F.R.S. With 56 Woodcuts. Fcp. 8vo., 6*s.*

HENDERSON.—PRACTICAL ELECTRICITY AND MAGNETISM. By JOHN HENDERSON, D.Sc. (Edin.), A.I.E.E. With 159 Illustrations and Diagrams. Crown 8vo., 6*s.* 6*d.*

JENKIN.—ELECTRICITY AND MAGNETISM. By FLEEMING JENKIN, F.R.S.S., L. and E., M.I.C.E. With 177 Illustrations. Fcp. 8vo., 3*s.* 6*d.*

JOUBERT.—ELEMENTARY TREATISE ON ELECTRICITY AND MAGNETISM. Founded on JOUBERT'S 'Traité Élémentairé d'Electricité'. By G. C. FOSTER, F.R.S., and E. ATKINSON, Ph.D. With 381 Illustrations. Crown 8vo., 7*s.* 6*d.*

JOYCE.—EXAMPLES IN ELECTRICAL ENGINEERING. By SAMUEL JOYCE, A.I.E.E. Crown 8vo., 5*s.*

LARDEN.—ELECTRICITY FOR PUBLIC SCHOOLS AND COLLEGES. By W. LARDEN, M.A. With 215 Illustrations. Cr. 8vo., 6*s.*

MERRIFIELD.—MAGNETISM AND DEVIATION OF THE COMPASS. By JOHN MERRIFIELD, LL.D., F.R.A.S., 18mo., 2*s.* 6*d.*

POYSER.—Works by A. W. POYSER, M.A.

MAGNETISM AND ELECTRICITY. With 235 Illustrations. Crown 8vo., 2*s.* 6*d.*

ADVANCED ELECTRICITY AND MAGNETISM. With 317 Illustrations. Crown 8vo., 4*s.* 6*d.*

SLINGO AND BROOKER.—Works by W. SLINGO and A. BROOKER.

ELECTRICAL ENGINEERING FOR ELECTRIC LIGHT ARTISANS AND STUDENTS. With 359 Illustrations. Crown 8vo., 12*s.*

PROBLEMS AND SOLUTIONS IN ELEMENTARY ELECTRICITY AND MAGNETISM. With 67 Illustrations. Cr. 8vo., 2*s.*

TYNDALL.—Works by JOHN TYNDALL, D.C.L., F.R.S. See p. 28.

TELEGRAPHY AND THE TELEPHONE.

HOPKINS.—THE TELEPHONE: Outlines of the Development of Transmitters and Receivers. By WILLIAM J. HOPKINS, Professor of Physics in the Drexel Institute, Philadelphia; Author of 'Telephone Lines and their Properties,' etc. With 7 Full-page Illustrations and 39 Diagrams. Crown 8vo., 3*s*. 6*d*.

PREECE AND SIVEWRIGHT.—TELEGRAPHY. By Sir W. H. PREECE, K.C.B., F.R.S., V.P.Inst., C.E., etc., Engineer-in-Chief and Electrician, Post Office Telegraphs; and Sir J. SIVEWRIGHT, K.C.M.G., General Manager, South African Telegraphs. With 267 Illustrations. Fcp. 8vo., 6*s*.

ENGINEERING, STRENGTH OF MATERIALS, ETC.

ANDERSON.—THE STRENGTH OF MATERIALS AND STRUCTURES: the Strength of Materials as depending on their Quality and as ascertained by Testing Apparatus. By Sir J. ANDERSON, C.E., LL.D., F.R.S.E. With 66 Woodcuts. Fcp. 8vo., 3*s*. 6*d*.

BARRY.—RAILWAY APPLIANCES: a Description of Details of Railway Construction subsequent to the completion of the Earthworks and Structures. By Sir JOHN WOLFE BARRY, K.C.B., F.R.S., M.I.C.E. With 218 Woodcuts. Fcp. 8vo., 4*s*. 6*d*.

GOODMAN.—MECHANICS APPLIED TO ENGINEERING. By JOHN GOODMAN, Wh.Sch., A.M.I.C.E., M.I.M.E., Professor of Engineering in the Yorkshire College, Leeds (Victoria University). With 620 Illustrations and numerous Examples. Crown 8vo., 7*s*. 6*d*. net.

LOW.—A POCKET-BOOK FOR MECHANICAL ENGINEERS. By DAVID ALLAN LOW (Whitworth Scholar), M.I.Mech.E., Professor of Engineering, East London Technical College (People's Palace), London. With over 1000 specially prepared Illustrations. Fcp. 8vo., gilt edges, rounded corners, 7*s*. 6*d*.

SMITH.—GRAPHICS, or the Art of Calculation by Drawing Lines, applied especially to Mechanical Engineering. By ROBERT H. SMITH, Professor of Engineering, Mason College, Birmingham. Part I. With separate Atlas of 29 Plates containing 97 Diagrams. 8vo., 15*s*.

STONEY.—THE THEORY OF STRESSES IN GIRDERS AND SIMILAR STRUCTURES; with Practical Observations on the Strength and other Properties of Materials. By BINDON B. STONEY, LL.D., F.R.S., M.I.C.E. With 5 Plates and 143 Illust. in the Text. Royal 8vo., 36*s*.

UNWIN.—Works by W CAWTHORNE UNWIN, F.R.S., B.S.C.

THE TESTING OF MATERIALS OF CONSTRUCTION. A Text-book for the Engineering Laboratory and a Collection of the Results of Experiment. With 5 Plates and 188 Illustrations and Diagrams in the Text. 8vo., 16*s*. net.

ON THE DEVELOPMENT AND TRANSMISSION OF POWER FROM CENTRAL STATIONS: being the Howard Lectures delivered at the Society of Arts in 1893. With 81 Diagrams. 8vo., 10*s*. net.

WARREN.—ENGINEERING CONSTRUCTION IN IRON, STEEL, AND TIMBER. By WILLIAM HENRY WARREN, Challis Professor of Civil and Mechanical Engineering, University of Sydney. With 13 Folding Plates and 375 Diagrams. Royal 8vo., 16*s*. net.

LONGMANS' CIVIL ENGINEERING SERIES.

Edited by the Author of 'Notes on Building Construction'.

TIDAL RIVERS: their (1) Hydraulics, (2) Improvement, (3) Navigation. By W. H. WHEELER, M.Inst.C.E. With 75 Illustrations. Medium 8vo., 16*s.* net.

NOTES ON DOCKS AND DOCK CONSTRUCTION. By C. COLSON, M.Inst.C.E., Deputy Civil Engineer-in-Chief, Admiralty. With 365 Illustrations. Medium 8vo., 21*s.* net.

PRINCIPLES AND PRACTICE OF HARBOUR CONSTRUCTION. By WILLIAM SHIELD, F.R.S.E., M.Inst.C.E., and Executive Engineer, National Harbour of Refuge, Peterhead, N.B. With 97 Illustrations. Medium 8vo., 15*s.* net.

CALCULATIONS IN HYDRAULIC ENGINEERING: a Practical Text-Book for the use of Students, Draughtsmen and Engineers. By T. CLAXTON FIDLER, M.Inst.C.E., Professor of Engineering, University College, Dundee.

Part I. Fluid Pressure and the Calculation of its Effects in Engineering Structures. With numerous Illustrations and Examples. 8vo., 6*s.* 6*d.* net.

RAILWAY CONSTRUCTION. By W. H. MILLS, M.I.C.E., Engineer-in-Chief of the Great Northern Railway of Ireland. With 516 Illustrations and Diagrams. 8vo., 18*s.* net.

NAVAL ARCHITECTURE.

ATTWOOD.—TEXT-BOOK OF THEORETICAL NAVAL ARCHITECTURE: a Manual for Students of Science Classes and Draughtsmen Engaged in Shipbuilders' and Naval Architects' Drawing Offices. By EDWARD LEWIS ATTWOOD, Assistant Constructor, Royal Navy; Member of the Institution of Naval Architects. With 114 Diagrams. Crown 8vo., 7*s.* 6*d.*

WATSON.—NAVAL ARCHITECTURE: A Manual of Laying-off Iron, Steel and Composite Vessels. By THOMAS H. WATSON, Lecturer on Naval Architecture at the Durham College of Science, Newcastle-upon-Tyne. With numerous Illustrations. Royal 8vo., 15*s.* net.

MACHINE DRAWING AND DESIGN.

LOW.—Works by DAVID ALLAN LOW, Professor of Engineering, East London Technical College (People's Palace).

IMPROVED DRAWING SCALES. 4*d.* in case.

AN INTRODUCTION TO MACHINE DRAWING AND DESIGN. With 153 Illustrations and Diagrams. Crown 8vo, 2*s.* 6*d.*

LOW AND BEVIS.—A MANUAL OF MACHINE DRAWING AND DESIGN. By DAVID ALLAN LOW and ALFRED WILLIAM BEVIS, M.I.Mech.E. With 700 Illustrations. 8vo., 7*s.* 6*d.*

UNWIN.—THE ELEMENTS OF MACHINE DESIGN. By W. CAWTHORNE UNWIN, F.R.S.

Part I. General Principles, Fastenings, and Transmissive Machinery. With 304 Diagrams, etc. Fcp. 8vo., 6*s.*

Part II. Chiefly on Engine Details. With 174 Woodcuts. Fcp. 8vo., 4*s.* 6*d.*

WORKSHOP APPLIANCES, ETC.

NORTHCOTT.—LATHES AND TURNING, Simple, Mechanical and Ornamental. By W. H. NORTHCOTT. With 338 Illustrations. 8vo., 18*s.*

SHELLEY.—WORKSHOP APPLIANCES, including Descriptions of some of the Gauging and Measuring Instruments, Hand-cutting Tools, Lathes, Drilling, Planeing, and other Machine Tools used by Engineers. By C. P. B. SHELLEY, M.I.C.E. With an additional Chapter on Milling by R. R. LISTER. With 323 Woodcuts. Fcp. 8vo., 5*s.*

MINERALOGY, METALLURGY, ETC.

BAUERMAN.—Works by HILARY BAUERMAN, F.G.S.

SYSTEMATIC MINERALOGY. With 373 Woodcuts and Diagrams. Fcp. 8vo., 6*s.*

DESCRIPTIVE MINERALOGY. With 236 Woodcuts and Diagrams. Fcp. 8vo., 6*s.*

GORE.—THE ART OF ELECTRO-METALLURGY, including all known Processes of Electro-Deposition. By G. GORE, LL.D., F.R.S. With 56 Woodcuts. Fcp. 8vo., 6*s.*

HUNTINGTON AND M'MILLAN.—METALS: their Properties and Treatment. By A. K. HUNTINGTON, Professor of Metallurgy in King's College, London, and W. G. M'MILLAN, Lecturer on Metallurgy in Mason's College, Birmingham. With 122 Illustrations. Fcp. 8vo., 7*s.* 6*d.*

RHEAD.—METALLURGY. An Elementary Text-Book. By E. C. RHEAD, Lecturer on Metallurgy at the Municipal Technical School, Manchester. With 94 Illustrations. Fcp. 8vo., 3*s.* 6*d.*

RUTLEY.—THE STUDY OF ROCKS: an Elementary Text-book of Petrology. By F. RUTLEY, F.G.S. With 6 Plates and 88 Woodcuts. Fcp. 8vo., 4*s.* 6*d.*

ASTRONOMY, NAVIGATION, ETC.

ABBOTT.—ELEMENTARY THEORY OF THE TIDES: the Fundamental Theorems Demonstrated without Mathematics and the Influence on the Length of the Day Discussed. By T. K. ABBOTT, B.D., Fellow and Tutor, Trinity College, Dublin. Crown 8vo., 2*s.*

BALL.—Works by Sir ROBERT S. BALL, LL.D., F.R.S.

ELEMENTS OF ASTRONOMY. With 130 Figures and Diagrams. Fcp. 8vo., 6*s.* 6*d.*

A CLASS-BOOK OF ASTRONOMY. With 41 Diagrams. Fcp. 8vo., 1*s.* 6*d.*

ASTRONOMY, NAVIGATION, ETC.—*Continued.*

CLERKE.—THE SYSTEM OF THE STARS. By AGNES M. CLERKE. With 6 Plates, and numerous Illustrations. 8vo., 21*s.*

DE CAMPIGNEULLES.—OBSERVATIONS TAKEN AT DUMRAON, BEHAR, INDIA, during the Eclipse of the 22nd January, 1898, by a Party of Jesuit Fathers of the Western Bengal Mission. By the Rev. V. DE CAMPIGNEULLES, S.J. With 14 Plates. 4to., 10*s.* 6*d.* net.

GILL.—TEXT-BOOK ON NAVIGATION AND NAUTICAL ASTRONOMY. By J. GILL, F.R.A.S., Head Master of the Liverpool Corporation Nautical College. 8vo., 10*s.* 6*d.*

HERSCHEL.—OUTLINES OF ASTRONOMY. By Sir JOHN F. W. HERSCHEL, Bart., K.H., etc. With 9 Plates and numerous Diagrams. 8vo., 12*s.*

LOWELL.—MARS. By PERCIVAL LOWELL, Fellow American Academy, Member Royal Asiatic Society, Great Britain and Ireland, etc. With 24 Plates. 8vo., 12*s.* 6*d.*

MARTIN.—NAVIGATION AND NAUTICAL ASTRONOMY. Compiled by Staff Commander W. R. MARTIN, R.N. Royal 8vo., 18*s.*

MERRIFIELD.—A TREATISE ON NAVIGATION. For the Use of Students. By J. MERRIFIELD, LL.D., F.R.A.S., F.M.S. With Charts and Diagrams. Crown 8vo., 5*s*

PARKER.—ELEMENTS OF ASTRONOMY. With Numerous Examples and Examination Papers. By GEORGE W. PARKER, M.A., of Trinity College, Dublin. With 84 Diagrams. 8vo., 5*s.* net.

WEBB.—CELESTIAL OBJECTS FOR COMMON TELESCOPES. By the Rev. T. W. WEBB, M.A., F.R.A.S. Fifth Edition, Revised and greatly Enlarged by the Rev. T. E. ESPIN, M.A., F.R.A.S. (Two Volumes.) Vol. I., with Portrait and a Reminiscence of the Author, 2 Plates, and numerous Illustrations. Crown 8vo., 6*s.* Vol. II., with numerous Illustrations. Crown 8vo., 6*s.* 6*d.*

WORKS BY RICHARD A. PROCTOR.

THE MOON: Her Motions, Aspect, Scenery, and Physical Condition. With many Plates and Charts, Wood Engravings, and 2 Lunar Photographs. Crown 8vo., 3*s.* 6*d.*

OTHER WORLDS THAN OURS: the Plurality of Worlds Studied Under the Light of Recent Scientific Researches. With 14 Illustrations; Map, Charts, etc. Crown 8vo., 3*s.* 6*d.*

OUR PLACE AMONG INFINITIES: a Series of Essays contrasting our Little Abode in Space and Time with the Infinities around us. Crown 8vo., 3*s.* 6*d.*

[OVER.

WORKS BY RICHARD A. PROCTOR—*Continued.*

MYTHS AND MARVELS OF ASTRONOMY. Crown 8vo., 3*s*. 6*d*.

LIGHT SCIENCE FOR LEISURE HOURS: Familiar Essays on Scientific Subjects, Natural Phenomena, etc. 3 vols., crown 8vo., 5*s*. each. Vol. I. Cheap Edition. Crown 8vo., 3*s*. 6*d*.

THE ORBS AROUND US; Essays on the Moon and Planets, Meteors and Comets, the Sun and Coloured Pairs of Suns. Crown 8vo., 3*s*. 6*d*.

THE EXPANSE OF HEAVEN: Essays on the Wonders of the Firmament. Crown 8vo., 3*s*. 6*d*.

OTHER SUNS THAN OURS: a Series of Essays on Suns—Old, Young, and Dead. With other Science Gleanings. Two Essays on Whist, and Correspondence with Sir John Herschel. With 9 Star-Maps and Diagrams. Crown 8vo., 3*s*. 6*d*.

HALF-HOURS WITH THE TELESCOPE: a Popular Guide to the Use of the Telescope as a means of Amusement and Instruction. With 7 Plates. Fcp. 8vo., 2*s*. 6*d*.

NEW STAR ATLAS FOR THE LIBRARY, the School, and the Observatory, in Twelve Circular Maps (with Two Index-Plates). With an Introduction on the Study of the Stars. Illustrated by 9 Diagrams. Cr. 8vo., 5*s*.

THE SOUTHERN SKIES: a Plain and Easy Guide to the Constellations of the Southern Hemisphere. Showing in 12 Maps the position of the principal Star-Groups night after night throughout the year. With an Introduction and a separate Explanation of each Map. True for every Year. 4to., 5*s*.

HALF-HOURS WITH THE STARS: a Plain and Easy Guide to the Knowledge of the Constellations. Showing in 12 Maps the position of the principal Star-Groups night after night throughout the year. With Introduction and a separate Explanation of each Map. True for every Year. 4to., 3*s*. 6*d*.

LARGER STAR ATLAS FOR OBSERVERS AND STUDENTS. In Twelve Circular Maps, showing 6000 Stars, 1500 Double Stars, Nebulæ, etc. With 2 Index-Plates. Folio, 15*s*.

THE STARS IN THEIR SEASONS: an Easy Guide to a Knowledge of the Star-Groups. In 12 Large Maps. Imperial 8vo., 5*s*.

ROUGH WAYS MADE SMOOTH. Familiar Essays on Scientific Subjects. Crown 8vo., 3*s*. 6*d*.

PLEASANT WAYS IN SCIENCE. Crown 8vo., 3*s*. 6*d*.

NATURE STUDIES. By R. A. PROCTOR, GRANT ALLEN, A. WILSON, T. FOSTER, and E. CLODD. Crown 8vo., 3*s*. 6*d*.

LEISURE READINGS. By R. A. PROCTOR, E. CLODD, A. WILSON, T. FOSTER, and A. C. RANYARD. Crown 8vo., 3*s*. 6*d*.

MANUFACTURES, TECHNOLOGY, ETC.

BELL.—JACQUARD WEAVING AND DESIGNING. By F. T. BELL, Medallist in Honours and Certificated Teacher in 'Linen Manufacturing' and in 'Weaving and Pattern Designing,' City and Guilds of London Institute. With 199 Diagrams. 8vo., 12*s.* net.

LUPTON.—MINING. An Elementary Treatise on the Getting of Minerals. By ARNOLD LUPTON, M.I.C.E., F.G.S., etc. With 596 Diagrams and Illustrations. Crown 8vo., 9*s.* net.

MORRIS AND WILKINSON.—THE ELEMENTS OF COTTON SPINNING. By JOHN MORRIS and F. WILKINSON. With a Preface by Sir B. A. DOBSON, C.E., M.I.M.E. With 169 Diagrams and Illustrations. Crown 8vo., 7*s.* 6*d.* net.

SHARP.—BICYCLES AND TRICYCLES: an Elementary Treatise on their Design and Construction. With Examples and Tables. By ARCHIBALD SHARP, B.Sc., Whitworth Scholar; Associate Member of the Institution of Civil Engineers. With 565 Illustrations and Diagrams. Crown 8vo., 15*s.*

TAYLOR.—COTTON WEAVING AND DESIGNING. By JOHN T. TAYLOR. With 373 Diagrams. Crown 8vo., 7*s.* 6*d.* net.

WATTS.—AN INTRODUCTORY MANUAL FOR SUGAR GROWERS. By FRANCIS WATTS, F.C.S., F.I.C. With 20 Illustrations. Crown 8vo., 6*s.*

PHYSIOGRAPHY AND GEOLOGY.

BIRD.—Works by CHARLES BIRD, B.A.

ELEMENTARY GEOLOGY. With Geological Map of the British Isles, and 247 Illustrations. Crown 8vo., 2*s.* 6*d.*

ADVANCED GEOLOGY. A Manual for Students in Advanced Classes and for General Readers. With over 300 Illustrations, a Geological Map of the British Isles (coloured), and a set of Questions for Examination. Crown 8vo., 7*s.* 6*d.*

GREEN.—PHYSICAL GEOLOGY FOR STUDENTS AND GENERAL READERS. By A. H. GREEN, M.A., F.G.S. With 236 Illustrations. 8vo., 21*s.*

MORGAN.—ELEMENTARY PHYSIOGRAPHY. Treated Experimentally. By ALEX. MORGAN, M.A., D.Sc., F.R.S.E., Lecturer in Mathematics and Science, Church of Scotland Training College, Edinburgh. With 4 Maps and 243 Diagrams. Crown 8vo., 2*s.* 6*d.*

THORNTON.—Works by J. THORNTON, M.A.

ELEMENTARY PRACTICAL PHYSIOGRAPHY (for Section I. of the New Syllabus of the Science and Art Department). With 215 Illustrations. Crown 8vo., 2*s.* 6*d.*

ELEMENTARY PHYSIOGRAPHY: an Introduction to the Study of Nature. With 12 Maps and 247 Illustrations. With Appendix on Astronomical Instruments and Measurements. Crown 8vo., 2*s.* 6*d.*

ADVANCED PHYSIOGRAPHY. With 6 Maps and 203 Illustrations. Crown 8vo., 4*s.* 6*d.*

HEALTH AND HYGIENE.

ASHBY.—HEALTH IN THE NURSERY. By HENRY ASHBY, M.D., F.R.C.P., Physician to the Manchester Children's Hospital, and Lecturer on the Diseases of Children at the Owens College. With 25 Illustrations. Crown 8vo., 3*s.* 6*d.*

BUCKTON. — HEALTH IN THE HOUSE; Twenty-five Lectures on Elementary Physiology. By Mrs. C. M. BUCKTON. With 41 Woodcuts and Diagrams. Crown 8vo., 2*s.*

CORFIELD.—THE LAWS OF HEALTH. By W. H. CORFIELD, M.A., M.D. Fcp. 8vo., 1*s.* 6*d.*

NOTTER AND FIRTH.—Works by J. L. NOTTER, M.A., M.D., and R. H. FIRTH, F.R.C.S.

HYGIENE. With 95 Illustrations. Crown 8vo., 3*s.* 6*d.*

PRACTICAL DOMESTIC HYGIENE. With 83 Illustrations. Crown 8vo., 2*s.* 6*d.*

POORE.—Works by GEORGE VIVIAN POORE, M.D.

ESSAYS ON RURAL HYGIENE. Crown 8vo., 6*s.* 6*d.*

THE DWELLING-HOUSE. With 36 Illustrations. Crown 8vo., 3*s.* 6*d.*

WILSON.—A MANUAL OF HEALTH-SCIENCE: adapted for use in Schools and Colleges. By ANDREW WILSON, F.R.S.E., F.L.S., etc. With 74 Illustrations. Crown 8vo., 2*s.* 6*d.*

NATURAL HISTORY.

BEDDARD.—THE STRUCTURE AND CLASSIFICATION OF BIRDS. By FRANK E. BEDDARD, M.A., F.R.S., Prosector and Vice-Secretary of the Zoological Society of London. With 252 Illustrations. 8vo., 21*s.* net.

FURNEAUX.—Works by WILLIAM FURNEAUX, F.R.G.S.

THE OUTDOOR WORLD; or, The Young Collector's Handbook. With 18 Plates, 16 of which are coloured, and 549 Illustrations in the Text. Crown 8vo., 7*s.* 6*d.*

LIFE IN PONDS AND STREAMS. With 8 Coloured Plates and 331 Illustrations in the Text. Crown 8vo., 7*s.* 6*d.*

BUTTERFLIES AND MOTHS (British). With 12 Coloured Plates and 241 Illustrations in the Text. 7*s.* 6*d.*

HUDSON.—BRITISH BIRDS. By W. H. HUDSON, C.M.Z.S. With 8 Coloured Plates from Original Drawings by A. THORBURN, and 8 Plates and 100 Figures by C. E. LODGE, and 3 Illustrations from Photographs. Crown 8vo., 7*s.* 6*d.*

STANLEY.—A FAMILIAR HISTORY OF BIRDS. By E. STANLEY, D.D., formerly Bishop of Norwich. With 160 Illustrations. Crown 8vo, 3*s.* 6*d.*

MEDICINE AND SURGERY.

ASHBY.—NOTES ON PHYSIOLOGY FOR THE USE OF STUDENTS PREPARING FOR EXAMINATION. By HENRY ASHBY, M.D. Lond., F.R.C.P., Physician to the General Hospital for Sick Children, Manchester; formerly Demonstrator of Physiology, Liverpool School of Medicine. Sixth Edition, thoroughly Revised. With 141 Illustrations. 18mo., 5*s.*

ASHBY AND WRIGHT.—THE DISEASES OF CHILDREN, MEDICAL AND SURGICAL. By HENRY ASHBY, M.D., Lond., F.R.C.P., Physician to the General Hospital for Sick Children, Manchester; and G. A. WRIGHT, B.A., M.B. Oxon., F.R.C.S., Eng., Assistant-Surgeon to the Manchester Royal Infirmary, and Surgeon to the Children's Hospital. Enlarged and Improved Edition. With 192 Illustrations. 8vo., 25*s.*

BENNETT.—Works by WILLIAM H. BENNETT, F.R.C.S., Surgeon to St. George's Hospital; Member of the Board of Examiners, Royal College of Surgeons of England.

CLINICAL LECTURES ON VARICOSE VEINS OF THE LOWER EXTREMITIES. With 3 Plates. 8vo., 6*s.*

ON VARICOCELE; A PRACTICAL TREATISE. With 4 Tables and a Diagram. 8vo., 5*s.*

CLINICAL LECTURES ON ABDOMINAL HERNIA: chiefly in relation to Treatment, including the Radical Cure. With 12 Diagrams in the Text. 8vo., 8*s.* 6*d.*

ON VARIX, ITS CAUSES AND TREATMENT, WITH ESPECIAL REFERENCE TO THROMBOSIS: an Address delivered at the Inaugural Meeting of the Nottingham Medico-Chirurgical Society, Session 1898-99. 8vo., 3*s.* 6*d.*

BENTLEY.—A TEXT-BOOK OF ORGANIC MATERIA MEDICA. Comprising a Description of the Vegetable and Animal Drugs of the British Pharmacopœia, with some others in common use. Arranged Systematically, and Especially Designed for Students. By ROBERT BENTLEY, M.R.C.S. Eng., F.L.S. With 62 Illustrations on Wood. Crown 8vo., 7*s.* 6*d.*

BRODIE.—THE ESSENTIALS OF EXPERIMENTAL PHYSIOLOGY. For the Use of Students. By T. G. BRODIE, M.D., Lecturer on Physiology, St. Thomas's Hospital Medical School. With 2 Plates and 177 Illustrations in the Text. Crown 8vo., 6*s.* 6*d.*

CABOT.—Works by RICHARD C. CABOT, M.D., Physician to Out-patients, Massachusetts General Hospital.

A GUIDE TO THE CLINICAL EXAMINATION OF THE BLOOD FOR DIAGNOSTIC PURPOSES. With 3 Coloured Plates and 28 Illustrations in the Text. 8vo., 16*s.*

THE SERUM DIAGNOSIS OF DISEASE. With 31 Temperature Charts and 9 Illustrations. Royal 8vo., 7*s.* 6*d.*

CHEYNE AND BURGHARD.—A MANUAL OF SURGIGAL TREATMENT. By W. WATSON CHEYNE, M.B., F.R.C.S., F.R.S., Professor of Surgery in King's College, London, Surgeon to King's College Hospital, etc.; and F. F. BURGHARD, M.D. and M.S., F.R.C.S., Teacher of Practical Surgery in King's College, London, Surgeon to King's College, Hospital (Lond.), etc. To be published in Six Parts.

Part I. The Treatment of General Surgical Diseases, including Inflammation, Suppuration, Ulceration, Gangrene, Wounds and their Complications, Infective Diseases and Tumours; the Administration of Anæsthetics. With 66 Illustrations. Royal 8vo., 10*s.* 6*d.*

MEDICINE AND SURGERY—*Continued.*

CLARKE.—Works by J. JACKSON CLARKE, M.B. Lond., F.R.C.S., Assistant Surgeon at the North-west London and City Orthopædic Hospitals, etc.

SURGICAL PATHOLOGY AND PRINCIPLES. With 194 Illustrations. Crown 8vo., 10*s.* 6*d.*

POST-MORTEM EXAMINATIONS IN MEDICO-LEGAL AND ORDINARY CASES. With Special Chapters on the Legal Aspects of Post-mortems, and on Certificates of Death. Fcp. 8vo., 2*s.* 6*d.*

COATS.—A MANUAL OF PATHOLOGY. By JOSEPH COATS, M.D., Professor of Pathology in the University of Glasgow. Third Edition. Revised throughout. With 507 Illustrations. 8vo., 31*s.* 6*d.*

COOKE.—Works by THOMAS COOK, F.R.C.S. Eng., B.A., B.Sc., M.D., Paris, Senior Assistant-Surgeon to the Westminster Hospital.

TABLETS OF ANATOMY. Being a Synopsis of Demonstrations given in the Westminster Hospital Medical School. Eleventh Edition in Three Parts, thoroughly brought up to date, and with over 700 Illustrations from all the best Sources, British and Foreign. Post 4to.

Part I. The Bones. 7*s.* 6*d.* net.

Part II. Limbs, Abdomen, Pelvis. 10*s.* 6*d.* net.

Part III. Head and Neck, Thorax, Brain. 10*s.* 6*d.* net.

APHORISMS IN APPLIED ANATOMY AND OPERATIVE SURGERY. Including 100 Typical *vivâ voce* Questions on Surface Marking, etc. Crown 8vo., 3*s.* 6*d.*

DISSECTION GUIDES. Aiming at Extending and Facilitating such Practical work in Anatomy as will be specially useful in connection with an ordinary Hospital Curriculum. 8vo., 10*s.* 6*d.*

DAKIN.—A HANDBOOK OF MIDWIFERY. By WILLIAM RADFORD DAKIN, M.D., F.R.C.P., Obstetric Physician and Lecturer on Midwifery at St. George's Hospital, etc. With 394 Illustrations. Large crown 8vo., 18*s.*

DICKINSON.—Works by W. HOWSHIP DICKINSON, M.D. Cantab., F.R.C.P., Physician to, and Lecturer on Medicine at, St. George's Hospital, Consulting Physician to the Hospital for Sick Children.

ON RENAL AND URINARY AFFECTIONS. With 12 Plates and 122 Woodcuts. Three Parts. 8vo., £3 4*s.* 6*d.*

THE TONGUE AS AN INDICATION OF DISEASE; being the Lumleian Lectures delivered at the Royal College of Physicians in March, 1888. 8vo., 7*s.* 6*d.*

OCCASIONAL PAPERS ON MEDICAL SUBJECTS, 1855-1896. 8vo., 12*s.*

MEDICINE AND SURGERY—*Continued.*

DUCKWORTH.—Works by SIR DYCE DUCKWORTH, M.D., LL.D., Fellow and Treasurer of the Royal College of Physicians, etc.

THE SEQUELS OF DISEASE: being the Lumleian Lectures delivered in the Royal College of Physicians, 1896. Together with Observations on Prognosis in Disease. 8vo., 10*s*. 6*d*.

THE INFLUENCE OF CHARACTER AND RIGHT JUDGMENT IN MEDICINE: the Harveian Oration, delivered before the Royal College of Physicians, 18th October, 1898. Post 4to., 2*s*. 6*d*.

ERICHSEN.—THE SCIENCE AND ART OF SURGERY; a Treatise on Surgical Injuries, Diseases, and Operations. By Sir JOHN ERIC ERICHSEN, Bart., F.R.S., LL.D. Edin., Hon. M.Ch. and F.R.C.S. Ireland, late Surgeon Extraordinary to H.M. the Queen. Illustrated by nearly 1000 Engravings on Wood. 2 vols. Royal 8vo., 48*s*.

FOWLER AND GODLEE.—THE DISEASES OF THE LUNGS. By JAMES KINGSTON FOWLER, M.A., M.D., F.R.C.P., Physician to the Middlesex Hospital and to the Hospital for Consumption and Diseases of the Chest, Brompton; late Examiner in Medicine at the University of Cambridge, and on the Conjoint Examining Board in England; and RICKMAN JOHN GODLEE, M.S., F.R.C.S., Fellow and Professor of Clinical Surgery, University College, London; Surgeon to University College Hospital and to the Hospital for Consumption and Diseases of the Chest, Brompton; Surgeon-in-Ordinary to Her Majesty's Household. With 160 Illustrations. 8vo., 25*s*.

GARROD.—Works by SIR ALFRED BARING GARROD, M.D., F.R.S., etc., Physician Extraordinary to H.M. the Queen; Consulting Physician to King's College Hospital; late Vice-President of the Royal College of Physicians.

A TREATISE ON GOUT AND RHEUMATIC GOUT (RHEUMATOID ARTHRITIS). Third Edition, thoroughly Revised and Enlarged. With 6 Plates, comprising 21 Figures (14 Coloured), and 27 Illustrations engraved on Wood. 8vo., 21*s*.

THE ESSENTIALS OF MATERIA MEDICA AND THERAPEUTICS. The Thirteenth Edition, Revised and Edited, under the supervision of the Author, by NESTOR TIRARD, M.D. Lond., F.R.C.P., Professor of Materia Medica and Therapeutics in King's College, London, etc. Crown 8vo., 12*s*. 6*d*.

GRAY.—ANATOMY, DESCRIPTIVE AND SURGICAL. By HENRY GRAY, F.R.S., late Lecturer on Anatomy at St. George's Hospital. The Fourteenth Edition, re-edited by T. PICKERING PICK, Surgeon to St. George's Hospital, Inspector of Anatomy in England and Wales, late Member of the Court of Examiners, Royal College of Surgeons of England. With 705 large Woodcut Illustrations, a large proportion of which are Coloured, the Arteries being coloured red, the Veins blue, and the Nerves yellow. The attachments of the muscles to the bones, in the section on Osteology, are also shown in coloured outline. Royal 8vo., 36*s*.

MEDICINE AND SURGERY—*Continued.*

HALLIBURTON.—Works by W. D. HALLIBURTON, M.D., F.R.S., F.R.C.P., Professor of Physiology in King's College, London.

A TEXT-BOOK OF CHEMICAL PHYSIOLOGY AND PATHOLOGY. With 104 Illustrations. 8vo., 28*s.*

ESSENTIALS OF CHEMICAL PHYSIOLOGY. With 77 Illustrations. 8vo., 5*s.*

LANG.—THE METHODICAL EXAMINATION OF THE EYE. Being Part I. of a Guide to the Practice of Ophthalmology for Students and Practitioners. By WILLIAM LANG, F.R.C.S. Eng., Surgeon to the Royal London Ophthalmic Hospital, Moorfields, etc. With 15 Illustrations. Crown 8vo., 3*s.* 6*d*

LIVEING.—HANDBOOK ON DISEASES OF THE SKIN. With especial reference to Diagnosis and Treatment. By ROBERT LIVEING, M.A. and M.D., Cantab., F.R.C.P. Lond., etc., Physician to the Department for Diseases of the Skin at the Middlesex Hospital, etc. Fcp. 8vo., 5*s.*

LUFF.—TEXT-BOOK OF FORENSIC MEDICINE AND TOXICOLOGY. By ARTHUR P. LUFF, M.D., B.Sc. (Lond.), Physician in Charge of Out-Patients and Lecturer on Medical Jurisprudence and Toxicology in St. Mary's Hospital. With 13 full-page Plates (1 in colours) and 33 Illustrations in the Text. 2 vols. Crown 8vo., 24*s.*

NEWMAN.—ON THE DISEASES OF THE KIDNEY AMENABLE TO SURGICAL TREATMENT. Lectures to Practitioners. By DAVID NEWMAN, M.D., Surgeon to the Western Infirmary Out-Door Department; Pathologist and Lecturer on Pathology at the Glasgow Royal Infirmary; Examiner in Pathology in the University of Glasgow; Vice-President, Glasgow Pathological and Clinical Society. 8vo., 8*s.*

PICK.—SURGERY: a Treatise for Students and Practitioners. By T. PICKERING PICK, Consulting Surgeon to St. George's Hospital; Senior Surgeon to the Victoria Hospital for Children; H.M. Inspector of Anatomy in England and Wales. With 441 Illustrations. Medium 8vo., 25*s.*

POOLE.—COOKERY FOR THE DIABETIC. By W. H. and Mrs. POOLE. With Preface by Dr. PAVY. Fcap. 8vo., 2*s.* 6*d.*

QUAIN.—A DICTIONARY OF MEDICINE; Including General Pathology, General Therapeutics, Hygiene, and the Diseases of Women and Children. By Various Writers. Edited by RICHARD QUAIN, Bart., M.D. Lond., LL.D. Edin. (Hon.) F.R.S., Physician Extraordinary to H.M. the Queen, etc. Assisted by FREDERICK THOMAS ROBERTS, M.D. Lond., B.Sc., Fellow of the Royal College of Physicians, Fellow of University College, etc.; and J. MITCHELL BRUCE, M.A. Abdn., M.D. Lond., Fellow of the Royal College of Physicians of London, etc. 2 Vols. Medium 8vo., 40*s.* net.

MEDICINE AND SURGERY—*Continued.*

QUAIN.—QUAIN'S (JONES) ELEMENTS OF ANATOMY. The Tenth Edition. Edited by EDWARD ALBERT SCHÄFER, F.R.S., Professor of Physiology in the University of Edinburgh; and GEORGE DANCER THANE, Professor of Anatomy in University College, London.

*** The several parts of this work form COMPLETE TEXT-BOOKS OF THEIR RESPECTIVE SUBJECTS.

VOL. I., PART I. EMBRYOLOGY. By E. A. SCHÄFER, F.R.S. With 200 Illustrations. Royal 8vo., 9*s.*

VOL. I., PART II. GENERAL ANATOMY OR HISTOLOGY. By E. A. SCHÄFER, F.R.S. With 291 Illustrations. Royal 8vo., 12*s.* 6*d.*

VOL. II., PART I. OSTEOLOGY—ARTHROLOGY. By G. D. THANE. With 224 Illus. Royal 8vo., 11*s.*

VOL. II., PART II. MYOLOGY—ANGEIOLOGY. By G. D. THANE. With 199 Illustrations. Royal 8vo., 16*s.*

VOL. III., PART I. THE SPINAL CORD AND BRAIN. By E. A. SCHÄFER, F.R.S. With 139 Illustrations. Royal 8vo., 12*s.* 6*d.*

VOL. III., PART II. THE NERVES. By G. D. THANE. With 102 Illustrations. Royal 8vo., 9*s.*

VOL. III., PART III. THE ORGANS OF THE SENSES. By E. A. SCHÄFER, F.R.S. With 178 Illustrations. Royal 8vo., 9*s.*

VOL. III., PART IV. SPLANCHNOLOGY. By E. A. SCHÄFER, F.R.S., and JOHNSON SYMINGTON, M.D. With 337 Illustrations. Royal 8vo., 16*s.*

APPENDIX. SUPERFICIAL AND SURGICAL ANATOMY. By Professor G. D. THANE and Professor R. J. GODLEE, M.S. With 29 Illustrations. Royal 8vo., 6*s.* 6*d.*

SCHÄFER.—THE ESSENTIALS OF HISTOLOGY. Descriptive and Practical. For the Use of Students. By E. A. SCHÄFER, F.R.S., Professor of Physiology in the University of Edinburgh; Editor of the Histological Portion of Quain's 'Anatomy'. Illustrated by nearly 400 Figures. Fifth Edition. Revised and Enlarged. 8vo., 8*s.* (Interleaved, 10*s.* 6*d.*)

SCHENK.—MANUAL OF BACTERIOLOGY. For Practitioners and Students. With especial reference to Practical Methods. By Dr. S. L. SCHENK, Professor (Extraordinary) in the University of Vienna. Translated from the German, with an Appendix, by W. R. DAWSON, B.A., M.D., Univ. Dub.; late University Travelling Prizeman in Medicine. With 100 Illustrations, some of which are coloured. 8vo., 10*s.* net.

SMALE AND COLYER. DISEASES AND INJURIES OF THE TEETH, including Pathology and Treatment: a Manual of Practical Dentistry for Students and Practitioners. By MORTON SMALE, M.R.C.S., L.S.A., L.D.S., Dental Surgeon to St. Mary's Hospital, Dean of the School, Dental Hospital of London, etc.; and J. F. COLYER, L.R.C.P., M.R.C.S., L.D.S., Assistant Dental Surgeon to Charing Cross Hospital, and Assistant Dental Surgeon to the Dental Hospital of London. With 334 Illustrations. Large crown 8vo., 15*s.*

SMITH (H. F.). THE HANDBOOK FOR MIDWIVES. By HENRY FLY SMITH, B.A., M.B. Oxon., M.R.C.S. 41 Woodcuts. Cr. 8vo., 5*s.*

STEVENSON.—WOUNDS IN WAR: the Mechanism of their Production and their Treatment. By Surgeon-Colonel W. F. STEVENSON (Army Medical Staff), A.B., M.B., M.Ch. Dublin University, Professor of Military Surgery, Army Medical School, Netley. With 86 Illustrations. 8vo., 18*s.*

TAPPEINER. — INTRODUCTION TO CHEMICAL METHODS OF CLINICAL DIAGNOSIS. By Dr. H. TAPPEINER, Professor of Pharmacology and Principal of the Pharmacological Institute of the University of Munich. Translated by EDMOND J. MCWEENEY, M.A., M.D. (Royal Univ. of Ireland), L.R.C.P.I., etc. Crown 8vo., 3*s.* 6*d.*

MEDICINE AND SURGERY—*Continued.*

TIRARD.—DIPHTHERIA AND ANTITOXIN. By NESTOR TIRARD, M.D. Lond., Fellow of the Royal College of Physicians; Fellow of King's College, London; Professor of Materia Medica and Therapeutics at King's College; Physician to King's College Hospital; and Senior Physician to the Evelina Hospital for Sick Children. 8vo., 7*s.* 6*d.*

WALLER.—Works by AUGUSTUS D. WALLER, M.D., Lecturer on Physiology at St. Mary's Hospital Medical School, London; late External Examiner at the Victorian University.

AN INTRODUCTION TO HUMAN PHYSIOLOGY. Third Edition, Revised. With 314 Illustrations. 8vo., 18*s.*

LECTURES ON PHYSIOLOGY. First Series. On Animal Electricity. 8vo., 5*s.* net.

EXERCISES IN PRACTICAL PHYSIOLOGY. Part I. Elementary Physiological Chemistry. By AUGUSTUS D. WALLER and W. LEGGE SYMES. 8vo., 1*s.* net. Part II. in the press. Part III. Physiology of the Nervous System; Electro-Physiology. 8vo., 2*s.* 6*d.* net.

WEICHSELBAUM.—THE ELEMENTS OF PATHOLOGICAL HISTOLOGY. With Special Reference to Practical Methods. By Dr. ANTON WEICHSELBAUM, Professor of Pathology in the University of Vienna. Translated by W. R. DAWSON, M.D. (Dub.), Demonstrator of Pathology in the Royal College of Surgeons, Ireland, late Medical Travelling Prizeman of Dublin University, etc. With 221 Figures, partly in Colours, a Cromo-lithographic Plate, and 7 Photographic Plates. Royal 8vo., 21*s.* net.

WILKS AND MOXON.—LECTURES ON PATHOLOGICAL ANATOMY. By Sir SAMUEL WILKS, Bart., M.D., F.R.S., President of the Royal College of Physicians, and Physician Extraordinary to H. M. the Queen, and the late WALTER MOXON, M.D., F.R.C.P., Physician to, and some time Lecturer on Pathology at, Guy's Hospital. Third Edition, thoroughly Revised. By Sir SAMUEL WILKS, Bart., M.D., LL.D., F.R.S. 8vo., 18*s.*

VETERINARY MEDICINE, ETC.

STEEL.—Works by JOHN HENRY STEEL, F.R.C.V.S., F.Z.S., A.V.D., late Professor of Veterinary Science and Principal of Bombay Veterinary College.

A TREATISE ON THE DISEASES OF THE DOG; being a Manual of Canine Pathology. Especially adapted for the use of Veterinary Practitioners and Students. With 88 Illustrations. 8vo., 10*s* 6*d.*

A TREATISE ON THE DISEASES OF THE OX; being a Manual of Bovine Pathology. Especially adapted for the use of Veterinary Practitioners and Students. With 2 Plates and 117 Woodcuts. 8vo. 15*s.*

A TREATISE ON THE DISEASES OF THE SHEEP; being a Manual of Ovine Pathology for the use of Veterinary Practitioners and Students. With Coloured Plate and 99 Woodcuts. 8vo., 12*s.*

OUTLINES OF EQUINE ANATOMY; a Manual for the use of Veterinary Students in the Dissecting Room. Crown 8vo., 7*s.* 6*d.*

FITZWYGRAM.—HORSES AND STABLES. By Major-General Sir F. FITZWYGRAM, Bart. With 56 pages of Illustrations. 8vo., 2*s.* 6*d.* net.

VETERINARY MEDICINE, ETC.—*Continued.*

SCHREINER.—THE ANGORA GOAT (published under the auspices of the South African Angora Goat Breeders' Association), and a Paper on the Ostrich (reprinted from the *Zoologist* for March, 1897). By S. C. CRONWRIGHT SCHREINER. With 26 Illustrations. 8vo., 10*s.* 6*d.*

'*STONEHENGE.*'—THE DOG IN HEALTH AND DISEASE. By 'STONEHENGE'. With 78 Wood Engravings. 8vo., 7*s.* 6*d.*

YOUATT.—Works by WILLIAM YOUATT.

THE HORSE. Revised and Enlarged by W. WATSON, M.R.C.V.S. With 52 Wood Engravings. 8vo., 7*s.* 6*d.*

THE DOG. Revised and Enlarged. With 33 Wood Engravings. 8vo., 6*s.*

PHYSIOLOGY, BIOLOGY, BACTERIOLOGY, AND ZOOLOGY.

(And see *MEDICINE AND SURGERY.*)

ASHBY.—NOTES ON PHYSIOLOGY. By HENRY ASHBY, M.D. With 141 Illustrations. Fcp. 8vo., 5*s.*

BARNETT.—THE MAKING OF THE BODY: a Children's Book on Anatomy and Physiology. By Mrs. S. A. BARNETT. With 113 Illustrations. Crown 8vo., 1*s.* 9*d.*

BEDDARD.—Works by FRANK E. BEDDARD, M.A. Oxon. F.R.S.; Prosector to the Zoological Society of London; Lecturer on Biology at Guy's Hospital; Examiner in Zoology and Comparative Anatomy in the University of London; lately Examiner in the Honours School of Morphology in the University of Oxford.

ELEMENTARY PRACTICAL ZOOLOGY. With 93 Illustrations. Crown 8vo., 2*s.* 6*d.*

THE STRUCTURE AND CLASSIFICATION OF BIRDS. With 252 Illustrations. 8vo., 21*s.* net.

BIDGOOD.—A COURSE OF PRACTICAL ELEMENTARY BIOLOGY. By JOHN BIDGOOD, B.Sc., F.L.S. With 226 Illustrations. Crown 8vo., 4*s.* 6*d.*

BRAY.—PHYSIOLOGY AND THE LAWS OF HEALTH, in Easy Lessons for Schools. By Mrs. CHARLES BRAY. Fcp. 8vo., 1*s.*

BRODIE.—THE ESSENTIALS OF EXPERIMENTAL PHYSIOLOGY. For the Use of Students. By T. G. BRODIE, M.D., Lecturer on Physiology, St. Thomas's Hospital Medical School. With 2 Plates and 177 Illustrations in the Text. Crown 8vo., 6*s.* 6*d.*

FRANKLAND.—MICRO-ORGANISMS IN WATER. Together with an Account of the Bacteriological Methods involved in their Investigation. Specially designed for the use of those connected with the Sanitary Aspects of Water-Supply. By PERCY FRANKLAND, Ph.D., B.Sc. (Lond.), F.R.S., and Mrs. PERCY FRANKLAND. With 2 Plates and Numerous Diagrams. 8vo., 16*s.* net.

FURNEAUX.—HUMAN PHYSIOLOGY. By W. FURNEAUX, F R G.S. With 218 Illustrations. Crown 8vo., 2*s.* 6*d.*

PHYSIOLOGY, BIOLOGY, BACTERIOLOGY, AND ZOOLOGY—*Cont.*

HUDSON AND GOSSE.—THE ROTIFERA, or 'WHEEL-ANIMACULES'. By C. T. HUDSON, LL.D., and P. H. GOSSE, F.R.S. With 30 Coloured and 4 Uncoloured Plates. In 6 Parts. 4to., 10*s.* 6*d.* each. Supplement 12*s.* 6*d.* Complete in 2 vols., with Supplement, 4to., £4 4*s.*

LEUMANN.—NOTES ON MICRO-ORGANISMS PATHOGENIC TO MAN. By Surgeon-Captain B. H. S. LEUMANN, M.B., Indian Medical Service. 8vo., 3*s.*

MACALISTER.—Works by ALEXANDER MACALISTER, M.D.

AN INTRODUCTION TO THE SYSTEMATIC ZOOLOGY AND MORPHOLOGY OF VERTEBRATE ANIMALS. With 41 Diagrams. 8vo., 10*s.* 6*d.*

ZOOLOGY OF THE INVERTEBRATE ANIMALS. With 59 Diagrams. Fcp. 8vo., 1*s.* 6*d.*

ZOOLOGY OF THE VERTEBRATE ANIMALS. With 77 Diagrams. Fcp. 8vo., 1*s.* 6*d.*

MOORE.—ELEMENTARY PHYSIOLOGY. By BENJAMIN MOORE, M.A., late Sharpey Research Scholar and Assistant Professor of Physiology at University College, London. With 125 Illustrations. Crown 8vo., 3*s.* 6*d.*

MORGAN.—ANIMAL BIOLOGY: an Elementary Text-Book. By C. LLOYD MORGAN, F.R.S., Principal of University College, Bristol. With 103 Illustrations. Crown 8vo., 8*s.* 6*d.*

SCHENK.—MANUAL OF BACTERIOLOGY, for Practitioners and Students, with Especial Reference to Practical Methods. By Dr. S. L. SCHENK. With 100 Illustrations, some Coloured. 8vo., 10*s.* net.

THORNTON.—HUMAN PHYSIOLOGY. By JOHN THORNTON, M.A. With 267 Illustrations, some Coloured. Crown 8vo., 6*s.*

BOTANY.

AITKEN.—ELEMENTARY TEXT-BOOK OF BOTANY. By EDITH AITKEN, late Scholar of Girton College. With 400 Diagrams. Crown 8vo., 4*s.* 6*d.*

BENNETT AND MURRAY.—HANDBOOK OF CRYPTOGAMIC BOTANY. By ALFRED W. BENNETT, M.A., B.Sc., F.L.S., Lecturer on Botany at St. Thomas's Hospital; and GEORGE MURRAY, F.L.S., Keeper of Botany, British Museum. With 378 Illustrations. 8vo., 16*s.*

CROSS AND BEVAN.—CELLULOSE: an Outline of the Chemistry of the Structural Elements of Plants. With Reference to their Natural History and Industrial Uses. By CROSS and BEVAN (C. F. Cross, E. J. Bevan, and C. Beadle). With 14 Plates. Crown 8vo., 12*s.* net.

CURTIS.—A TEXT-BOOK OF GENERAL BOTANY. By CARLTON C. CURTIS, A.M., Ph.D., Tutor in Botany in Columbia University, U.S.A. With 87 Illustrations. 8vo., 12*s.* net.

DE TABLEY.—THE FLORA OF CHESHIRE. By the late LORD DE TABLEY (Hon. J. BYRNE LEICESTER WARREN, M.A.). Edited by SPENCER MOORE. With a Biographical Notice of the Author by Sir MOUNTSTUART GRANT DUFF. With a Map of Cheshire and a Photogravure Portrait. Crown 8vo., 10*s.* 6*d.* net.

BOTANY—*Continued.*

EDMONDS.—Works by HENRY EDMONDS, B.Sc., London.

ELEMENTARY BOTANY. With 319 Illustrations. Cr.8vo., 2*s*. 6*d*.

BOTANY FOR BEGINNERS. With 85 Illustrations. Fcp. 8vo., 1*s*. 6*d*.

FARMER.—A PRACTICAL INTRODUCTION TO THE STUDY OF BOTANY: Flowering Plants. By J. BRETLAND FARMER, M.A., Professor of Botany in the Royal College of Science, London; formerly Fellow of Magdalen College, Oxford. With 121 Illustrations. Crown 8vo, 2*s*. 6*d*.

KITCHENER.—A YEAR'S BOTANY. Adapted to Home and School Use. By FRANCES A. KITCHENER. With 195 Illustrations. Cr. 8vo., 5*s*.

LINDLEY AND MOORE.—THE TREASURY OF BOTANY. Edited by J. LINDLEY, M.D., F.R.S., and T. MOORE, F.L.S. With 20 Steel Plates and numerous Woodcuts. Two parts. Fcp. 8vo., 12*s*.

McNAB.—CLASS-BOOK OF BOTANY. By W. R. McNAB.

MORPHOLOGY AND PHYSIOLOGY. With 42 Diagrams. Fcp. 8vo., 1*s*. 6*d*.

CLASSIFICATION OF PLANTS. With 118 Diagrams. Fcp. 8vo., 1*s*. 6*d*.

SORAUER.—A POPULAR TREATISE ON THE PHYSIOLOGY OF PLANTS. By Dr. PAUL SORAUER. Translated by F. E. WEISS, B.Sc., F.L.S. With 33 Illustrations. 8vo., 9*s*. net.

THOMÉ AND BENNETT.—STRUCTURAL AND PHYSIOLOGICAL BOTANY. By OTTO WILHELM THOMÉ and by ALFRED W. BENNETT, B.Sc., F.L.S. With Coloured Map and 600 Woodcuts. Fcp. 8vo., 6*s*.

TUBEUF.—DISEASES OF PLANTS INDUCED BY CRYPTOGAMIC PARASITES. Introduction to the Study of Pathogenic Fungi, Slime Fungi, Bacteria and Algæ. By Dr. KARL FREIHERR VON TUBEUF, Privatdocent in the University of Munich. English Edition by WILLIAM G. SMITH, B.Sc., Ph.D., Lecturer on Plant Physiology, University of Edinburgh. With 330 Illustrations. Royal 8vo., 18*s*. net.

WATTS.—A SCHOOL FLORA. For the use of Elementary Botanical Classes. By W. MARSHALL WATTS, D.Sc. Lond. Cr, 8vo., 2*s*. 6*d*.

AGRICULTURE.

ADDYMAN.—AGRICULTURAL ANALYSIS. A Manual of Quantitative Analysis for Students of Agriculture. By FRANK T. ADDYMAN, B.Sc. (Lond.), F.I.C. With 49 Illustrations. Crown 8vo., 5*s*. net.

COLEMAN AND ADDYMAN.—PRACTICAL AGRICULTURAL CHEMISTRY. By J. BERNARD COLEMAN, A.R.C.Sc., F.I.C., and FRANK T. ADDYMAN, B.Sc. (Lond.), F.I.C. With 24 Illustrations. Crown 8vo., 1*s* 6*d*. net.

COOKE.—THE FOUNDATIONS OF SCIENTIFIC AGRICULTURE. By SAMUEL COOKE, A.M., A.M.I.C.E., F.I.C., F.G.S., Cor.M.R.H.S. With 85 Illustrations and Examination Questions. Cr.8vo., 4*s*. 6*d*.

WEBB.—Works by HENRY J. WEBB, Ph.D., B.Sc. (Lond.).

ELEMENTARY AGRICULTURE. A Text-Book specially adapted to the requirements of the Science and Art Department, the Junior Examination of the Royal Agricultural Society, and other Elementary Examinations. With 34 Illustrations. Crown 8vo., 2*s*. 6*d*.

AGRICULTURE. A Manual for Advanced Science Students. With 100 Illustrations. Crown 8vo., 7*s*. 6*d*. net.

WORKS BY JOHN TYNDALL, D.C.L., LL.D., F.R.S.

FRAGMENTS OF SCIENCE: a Series of Detached Essays, Addresses, and Reviews. 2 vols. Crown 8vo., 16*s.*

Vol. I.—The Constitution of Nature—Radiation—On Radiant Heat in Relation to the Colour and Chemical Constitution of Bodies—New Chemical Reactions produced by Light—On Dust and Disease—Voyage to Algeria to observe the Eclipse—Niagara—The Parallel Roads of Glen Roy—Alpine Sculpture—Recent Experiments on Fog-Signals—On the Study of Physics—On Crystalline and Slaty Cleavage—On Paramagnetic and Diamagnetic Forces—Physical Basis of Solar Chemistry—Elementary Magnetism—On Force—Contributions to Molecular Physics—Life and Letters of FARADAY—The Copley Medallist of 1870—The Copley Medallist of 1871—Death by Lightning—Science and the Spirits.

Vol. II.—Reflections on Prayer and Natural Law—Miracles and Special Providences—On Prayer as a Form of Physical Energy—Vitality—Matter and Force—Scientific Materialism—An Address to Students—Scientific Use of the Imagination—The Belfast Address—Apology for the Belfast Address—The Rev. JAMES MARTINEAU and the Belfast Address—Fermentation, and its Bearings on Surgery and Medicine—Spontaneous Generation—Science and Man—Professor VIRCHOW and Evolution—The Electric Light.

NEW FRAGMENTS. Crown 8vo., 10*s.* 6*d.*

CONTENTS.—The Sabbath—Goethe's 'Farbenlehre'—Atoms, Molecules, and Ether Waves—Count Rumford—Louis Pasteur, his Life and Labours—The Rainbow and its Congeners—Address delivered at the Birkbeck Institution on October 22, 1884—Thomas Young—Life in the Alps—About Common Water—Personal Recollections of Thomas Carlyle—On Unveiling the Statue of Thomas Carlyle—On the Origin, Propagation, and Prevention of Phthisis—Old Alpine Jottings—A Morning on Alp Lusgen.

LECTURES ON SOUND. With Frontispiece of Fog-Syren, and 203 other Woodcuts and Diagrams in the Text. Crown 8vo., 10*s.* 6*d.*

HEAT, A MODE OF MOTION. With 125 Woodcuts and Diagrams. Crown 8vo., 12*s.*

LECTURES ON LIGHT DELIVERED IN THE UNITED STATES IN 1872 AND 1873. With Portrait, Lithographic Plate, and 59 Diagrams. Crown 8vo., 5*s.*

ESSAYS ON THE FLOATING MATTER OF THE AIR IN RELATION TO PUTREFACTION AND INFECTION. With 24 Woodcuts. Crown 8vo., 7*s.* 6*d.*

RESEARCHES ON DIAMAGNETISM AND MAGNECRYSTALLIC ACTION; including the Question of Diamagnetic Polarity. Crown 8vo., 12*s.*

NOTES OF A COURSE OF NINE LECTURES ON LIGHT. delivered at the Royal Institution of Great Britain, 1869. Crown 8vo., 1*s.* 6*d.*

NOTES OF A COURSE OF SEVEN LECTURES ON ELECTRICAL PHENOMENA AND THEORIES, delivered at the Royal Institution of Great Britain, 1870. Crown 8vo., 1*s.* 6*d.*

LESSONS IN ELECTRICITY AT THE ROYAL INSTITUTION 1875-1876. With 58 Woodcuts and Diagrams. Crown 8vo., 2*s.* 6*d.*

THE GLACIERS OF THE ALPS: being a Narrative of Excursions and Ascents. An Account of the Origin and Phenomena of Glaciers, and an Exposition of the Physical Principles to which they are related. With numerous Illustrations. Crown 8vo., 6*s.* 6*d.* net.

HOURS OF EXERCISE IN THE ALPS. With 7 Illustrations. Crown 8vo., 6*s.* 6*d.* net.

FARADAY AS A DISCOVERER. Crown 8vo., 3*s.* 6*d.*

TEXT-BOOKS OF SCIENCE.

PHOTOGRAPHY. By Captain W. DE WIVELESLIE ABNEY, C.B., F.R.S. With 105 Illustrations. Fcp. 8vo., 3*s*. 6*d*.

THE STRENGTH OF MATERIAL AND STRUCTURES. By Sir J. ANDERSON, C.E. With 66 Illustrations. Fcp. 8vo., 3*s*. 6*d*.

RAILWAY APPLIANCES. By Sir JOHN WOLFE BARRY, K.C.B., F.R.S., M.I.C.E. With 218 Illustrations. Fcp. 8vo., 4*s*. 6*d*.

INTRODUCTION TO THE STUDY OF INORGANIC CHEMISTRY. By WILLIAM ALLEN MILLER, M.D., LL.D., F.R.S. With 72 Illustrations. 3*s*. 6*d*.

QUANTITATIVE CHEMICAL ANALYSIS. By T. E. THORPE, F.R.S., Ph.D. With 88 Illustrations. Fcp. 8vo., 4*s*. 6*d*.

QUALITATIVE ANALYSIS AND LABORATORY PRACTICE. By T. E. THORPE, Ph.D., F.R.S., and M. M. PATTISON MUIR, M.A. and F.R.S.E. With Plate of Spectra and 57 Illustrations. Fcp. 8vo., 3*s*. 6*d*.

INTRODUCTION TO THE STUDY OF CHEMICAL PHILOSOPHY. By WILLIAM A. TILDEN, D.Sc., London, F.R.S. With 5 Illustrations. With or without Answers to Problems. Fcp. 8vo., 4*s*. 6*d*

ELEMENTS OF ASTRONOMY. By Sir R. S. BALL, LL.D., F.R.S. With 130 Illustrations. Fcp. 8vo., 6*s*. 6*d*.

SYSTEMATIC MINERALOGY. By HILARY BAUERMAN, F.G.S. With 373 Illustrations. Fcp. 8vo., 6*s*.

DESCRIPTIVE MINERALOGY. By HILARY BAUERMAN, F.G.S., etc. With 236 Illustrations. Fcp. 8vo., 6*s*.

METALS: THEIR PROPERTIES AND TREATMENT. By A. K. HUNTINGTON and W. G. MCMILLAN. With 122 Illustrations. Fcp. 8vo., 7*s*. 6*d*.

THEORY OF HEAT. By J. CLERK MAXWELL, M.A., LL.D., Edin., F.R.SS., L. & E. With 38 Illustrations. Fcp. 8vo., 4*s*. 6*d*.

PRACTICAL PHYSICS. By R. T. GLAZEBROOK, M.A., F.R.S., and W. N. SHAW, M.A. With 134 Illustrations. Fcp. 8vo., 7*s*. 6*d*.

PRELIMINARY SURVEY AND ESTIMATES. By THEODORE GRAHAM GRIBBLE, Civil Engineer. Including Elementary Astronomy, Route Surveying, Tacheometry, Curve-ranging, Graphic Mensuration, Estimates, Hydrography and Instruments. With 133 Illustrations. Fcp. 8vo., 7*s*. 6*d*.

ALGEBRA AND TRIGONOMETRY. By WILLIAM NATHANIEL GRIFFIN, B.D. 3*s*. 6*d*. Notes on, with Solutions of the more difficult Questions. Fcp. 8vo., 3*s*. 6*d*.

THE STEAM ENGINE. By GEORGE C. V. HOLMES, Secretary of the Institution of Naval Architects. With 212 Illustrations. Fcp. 8vo., 6*s*.

ELECTRICITY AND MAGNETISM. By FLEEMING JENKIN, F.R.SS., L. & E. With 177 Illustrations. Fcp. 8vo., 3*s*. 6*d*.

THE ART OF ELECTRO-METALLURGY. By G. GORE, LL.D., F.R.S. With 56 Illus. Fcp. 8vo., 6*s*.

TELEGRAPHY. By Sir W. H. PREECE, K.C.B., F.R.S., M.I.C.E., and Sir J. SIVEWRIGHT, M.A., K.C.M.G. With 267 Illustrations. Fcp. 8vo., 6*s*.

PHYSICAL OPTICS. By R. T. GLAZEBROOK, M.A., F.R.S. With 183 Illustrations. Fcp. 8vo., 6*s*.

TECHNICAL ARITHMETIC AND MENSURATION. By CHARLES W. MERRIEFIELD, F.R.S. 3*s*. 6*d*. Key, by the Rev. JOHN HUNTER, M.A. Fcp. 8vo., 3*s*. 6*d*.

THE STUDY OF ROCKS. By FRANK RUTLEY, F.G.S. With 6 Plates and 88 Illustrations Fcp. 8vo., 4*s*. 6*d*.

WORKSHOP APPLIANCES, including Descriptions of some of the Machine Tools used by Engineers. By C. P. B. SHELLEY, M.I.C.E. With 323 Illustrations. Fcp. 8vo., 5*s*.

ELEMENTS OF MACHINE DESIGN. By W. CAWTHORNE UNWIN, F.R.S., B.Sc., M.I.C.E.

PART I. General Principles, Fastenings and Transmissive Machinery. With 304 Illustrations. 6*s*.

PART II. Chiefly on Engine Details. With 174 Illustrations. Fcp. 8vo., 4*s*. 6*d*.

STRUCTURAL AND PHYSIOLOGICAL BOTANY. By OTTO WILHELM THOMÉ, and A. W. BENNETT, M.A., B.Sc., F.L.S. With 600 Illustrations. Fcp. 8vo., 6*s*.

PLANE AND SOLID GEOMETRY. By H. W. WATSON, M.A. Fcp. 8vo. 3*s*. 6*d*.

ADVANCED SCIENCE MANUALS.

BUILDING CONSTRUCTION. By the Author of 'Rivington's Notes on Building Construction'. With 385 Illustrations and an Appendix of Examination Questions. Crown 8vo., 4*s*. 6*d*.

THEORETICAL MECHANICS. Solids, including Kinematics, Statics, and Kinetics. By A. THORNTON, M.A., F.R.A.S. With 220 Illustrations, 130 Worked Examples, and over 900 Examples from Examination Papers, etc. Crown 8vo., 4*s*. 6*d*.

HEAT. By MARK R. WRIGHT, Hon. Inter. B.Sc. (Lond.). With 136 Illustrations and numerous Examples and Examination Papers. Crown 8vo., 4*s*. 6*d*.

LIGHT. By W. J. A. EMTAGE, M.A. With 232 Illustrations. Cr. 8vo., 6*s*.

MAGNETISM AND ELECTRICITY. By ARTHUR WILLIAM POYSER, M.A. With 317 Illustrations. Crown 8vo., 4*s*. 6*d*.

INORGANIC CHEMISTRY, THEORETICAL AND PRACTICAL. A Manual for Students in Advanced Classes of the Science and Art Department. By WILLIAM JAGO, F.C.S., F.I.C. With Plate of Spectra and 78 Woodcuts. Crown 8vo., 4*s*. 6*d*.

GEOLOGY: a Manual for Students in Advanced Classes and for General Readers. By CHARLES BIRD, B.A. (Lond.), F.G.S. With over 300 Illustrations, a Geological Map of the British Isles (coloured), and a set of Questions for Examination. Crown 8vo., 7*s*. 6*d*.

HUMAN PHYSIOLOGY: a Manual for Students in advanced Classes of the Science and Art Department. By JOHN THORNTON, M.A. With 268 Illustrations, some of which are Coloured, and a set of Questions for Examination. Crown 8vo., 6*s*.

PHYSIOGRAPHY. By JOHN THORNTON, M.A. With 6 Maps, 203 Illustrations, and Coloured Plate of Spectra. Crown 8vo., 4*s*. 6*d*.

AGRICULTURE. By HENRY J. WEBB, Ph.D., B.Sc. With 100 Illustrations. Crown 8vo., 7*s*. 6*d*. net.

HYGIENE. By J. LANE NOTTER, M.A., M.D., Professor of Hygiene in the Army Medical School, Netley, Surgeon-Colonel, Army Medical Staff; and R. H. FIRTH, F.R.C.S., Assistant Professor of Hygiene in the Army Medical School, Netley, Surgeon-Major, Army Medical Staff. With 95 Illustrations. Crown 8vo., 3*s*. 6*d*.

ELEMENTARY SCIENCE MANUALS.

*** *Written specially to meet the requirements of the ELEMENTARY STAGE OF SCIENCE SUBJECTS as laid down in the Syllabus of the Directory of the SCIENCE AND ART DEPARTMENT.*

PRACTICAL, PLANE, AND SOLID GEOMETRY, including Graphic Arithmetic. By I. H. MORRIS. Fully Illustrated with Drawings. Crown 8vo., 2*s*. 6*d*.

GEOMETRICAL DRAWING FOR ART STUDENTS. Embracing Plane Geometry and its Applications, the Use of Scales, and the Plans and Elevations of Solids, as required for the Examinations of the Science and Art Department. By I. H. MORRIS. Crown 8vo., 1*s*. 6*d*.

TEXT-BOOK ON PRACTICAL, SOLID, OR DESCRIPTIVE GEOMETRY. By DAVID ALLAN LOW (Whitworth Scholar). Part I. Crown 8vo., 2*s*. Part II. Crown 8vo., 3*s*.

AN INTRODUCTION TO MACHINE DRAWING AND DESIGN. By DAVID ALLAN LOW. With 153 Illustrations. Crown 8vo., 2*s*. 6*d*.

ELEMENTARY SCIENCE MANUALS—*Continued.*

BUILDING CONSTRUCTION AND DRAWING. By EDWARD J. BURRELL. With 308 Illustrations and Working Drawings. Crown 8vo., 2*s.* 6*d.*

AN ELEMENTARY COURSE OF MATHEMATICS. Containing Arithmetic; Euclid (Book I., with Deductions and Exercises); and Algebra. Crown 8vo., 2*s.* 6*d.*

THEORETICAL MECHANICS. Including Hydrostatics and Pneumatics. By J. E. TAYLOR, M.A., B.Sc. With numerous Examples and Answers, and 175 Diagrams and Illustrations. Crown 8vo., 2*s.* 6*d.*

THEORETICAL MECHANICS—SOLIDS. By J. E. TAYLOR, M.A., B.Sc. (Lond.). With 163 Illustrations, 120 Worked Examples, and over 500 Examples from Examination Papers, etc. Crown 8vo., 2*s.* 6*d.*

THEORETICAL MECHANICS—FLUIDS. By J. E. TAYLOR, M.A., B.Sc. (Lond.). With 122 Illustrations, numerous Worked Examples, and about 500 Examples from Examination Papers, etc. Crown 8vo., 2*s.* 6*d.*

A MANUAL OF MECHANICS. With 138 Illustrations and Diagrams, and 188 Examples taken from Examination Papers, with Answers. By T. M. GOODEVE, M.A. Crown 8vo., 2*s.* 6*d.*

SOUND, LIGHT, AND HEAT. By MARK R. WRIGHT, M.A. With 160 Diagrams and Illustrations. Crown 8vo., 2*s.* 6*d.*

METALLURGY: an Elementary Text-Book. By E. L. RHEAD. With 94 Illustrations. Crown 8vo., 3*s.* 6*d.*

PHYSICS. Alternative Course. By MARK R. WRIGHT, M.A. With 242 Illustrations. Crown 8vo., 2*s.* 6*d.*

MAGNETISM AND ELECTRICITY. By A. W. POYSER, M.A. With 235 Illustrations. Crown 8vo., 2*s.* 6*d.*

ORGANIC CHEMISTRY: the Fatty Compounds. By R. LLOYD WHITELEY, F.I.C., F.C.S. With 45 Illustrations. Crown 8vo., 3*s.* 6*d.*

INORGANIC CHEMISTRY, THEORETICAL AND PRACTICAL. By WILLIAM JAGO, F.C.S., F.I.C. With 63 Illustrations and numerous Questions and Exercises. Fcp. 8vo., 2*s.* 6*d.*

AN INTRODUCTION TO PRACTICAL INORGANIC CHEMISTRY. By WILLIAM JAGO, F.C.S., F.I.C. Crown 8vo., 1*s.* 6*d.*

PRACTICAL CHEMISTRY: the Principles of Qualitative Analysis. By WILLIAM A. TILDEN, D.Sc. Fcp. 8vo., 1*s.* 6*d.*

ELEMENTARY INORGANIC CHEMISTRY. By WILLIAM FURNEAUX, F.R.G.S. Crown 8vo., 2*s.* 6*d.*

ELEMENTARY GEOLOGY. By CHARLES BIRD, B.A., F.G.S. With Coloured Geological Map of the British Islands, and 247 Illustrations. Crown 8vo., 2*s.* 6*d.*

HUMAN PHYSIOLOGY. By WILLIAM FURNEAUX, F.R.G.S. With 218 Illustrations. Crown 8vo., 2*s.* 6*d.*

A COURSE OF PRACTICAL ELEMENTARY BIOLOGY. By J. BIDGOOD, B.Sc. With 226 Illustrations. Crown 8vo., 4*s.* 6*d.*

ELEMENTARY BOTANY, THEORETICAL AND PRACTICAL. By HENRY EDMONDS, B.Sc. With 342 Illustrations. Crown 8vo, 2*s.* 6*d.*

STEAM. By WILLIAM RIPPER, Member of the Institution of Mechanical Engineers. With 142 Illustrations. Crown 8vo., 2*s.* 6*d.*

ELEMENTARY PHYSIOGRAPHY. By J. THORNTON, M.A. With 12 Maps and 249 Illustrations. With Appendix on Astronomical Instruments and Measurements. Crown 8vo., 2*s.* 6*d.*

AGRICULTURE. By HENRY J. WEBB, Ph.D. With 34 Illustrations. Crown 8vo., 2*s.* 6*d.*

THE LONDON SCIENCE CLASS-BOOKS.

Edited by G. CAREY FOSTER, F.R.S., and by Sir PHILIP MAGNUS, B.Sc., B.A. of the City and Guilds of London Institute.

ASTRONOMY. By Sir ROBERT STAWELL BALL, LL.D., F.R.S. With 41 Diagrams. Fcp. 8vo., 1*s.* 6*d.*

MECHANICS. By SIR ROBERT STAWELL BALL, LL.D., F.R.S. With 89 Diagrams. Fcp. 8vo., 1*s.* 6*d.*

THE LAWS OF HEALTH. By W. H. CORFIELD, M A., M.D., F.R.C.P. With 22 Illustrations. Fcp. 8vo., 1*s.* 6*d.*

MOLECULAR PHYSICS AND SOUND. By FREDERICK GUTHRIE, F.R.S. With 91 Diagrams. Fcp. 8vo., 1*s.* 6*d.*

GEOMETRY, CONGRUENT FIGURES. By O. HENRICI, Ph.D., F.R.S. With 141 Diagrams. Fcp. 8vo., 1*s.* 6*d.*

ZOOLOGY OF THE INVERTEBRATE ANIMALS. By ALEXANDER MACALISTER, M.D. With 59 Diagrams. Fcp. 8vo., 1*s.* 6*d.*

ZOOLOGY OF THE VERTEBRATE ANIMALS. By ALEXANDER MACALISTER, M.D. With 77 Diagrams Fcp. 8vo., 1*s.* 6*d.*

HYDROSTATICS AND PNEUMATICS. By Sir PHILIP MAGNUS, B.Sc., B.A. With 79 Diagrams. Fcp. 8vo., 1*s.* 6*d.* (To be had also *with Answers*, 2*s.*) The Worked Solutions of the Problems. 2*s.*

BOTANY. Outlines of the Classification of Plants. By W. R. MCNAB, M.D. With 118 Diagrams. Fcp. 8vo., 1*s.* 6*d.*

BOTANY. Outlines of Morphology and Physiology. By W. R. MCNAB, M.D. With 42 Diagrams. Fcp. 8vo., 1*s.* 6*d.*

THERMODYNAMICS. By RICHARD WORMELL, M.A., D.Sc. With 41 Diagrams. Fcp. 8vo., 1*s.* 6*d.*

PRACTICAL ELEMENTARY SCIENCE SERIES.

ELEMENTARY PRACTICAL PHYSIOGRAPHY. (Section I.) By JOHN THORNTON, M.A., Head Master of the Central Higher Grade School, Bolton. With 215 Illustrations and a Coloured Spectrum. Crown 8vo., 2*s.* 6*d.*

PRACTICAL DOMESTIC HYGIENE. By J. LANE NOTTER, M.A., M.D., Professor of Hygiene in the Army Medical School, Netley, etc.; and R. H. FIRTH, F.R.C.S., Assistant Professor of Hygiene in the Army Medical School, Netley, etc. With 83 Illustrations. Crown 8vo., 2*s.* 6*d.*

A PRACTICAL INTRODUCTION TO THE STUDY OF BOTANY: Flowering Plants. By J. BRETLAND FARMER, M.A., Professor of Botany in the Royal College of Science, London; formerly Fellow of Magdalen College, Oxford. With 121 Illustrations. Crown 8vo, 2*s.* 6*d.*

ELEMENTARY PRACTICAL CHEMISTRY. By G. S. NEWTH, F.I.C., F.C.S., Demonstrator in the Royal College of Science, London, etc. With 108 Illustrations and 254 Experiments. Crown 8vo., 2*s.* 6*d.*

ELEMENTARY PRACTICAL PHYSICS. By W. WATSON, B.Sc., Assistant Professor in Physics in the Royal College of Science, London, etc. With 120 Illustrations and 193 Exercises. Crown 8vo., 2*s.* 6*d.*

ELEMENTARY PRACTICAL ZOOLOGY. By FRANK E. BEDDARD, M.A. Oxon., F.R.S., Prosector to the Zoological Society of London; Lecturer on Biology at Guy's Hospital. With 93 Illustrations. Crown 8vo., 2*s.* 6*d.*

OTHER VOLUMES IN PREPARATION.

10,000/9/99.